Review

New Phytologist

Tansley review

Plant hemoglobins: a journey from unicellular green algae to vascular plants

Author for correspondence:
Manuel Becana
Tel: +34 976 716055
Email: becana@eead.csic.es

Received: *12 November 2019*
Accepted: *24 December 2019*

Manuel Becana[1], Inmaculada Yruela[1,2], Gautam Sarath[3], Pilar Catalán[2,4] and Mark S. Hargrove[5]

[1]Departamento de Nutrición Vegetal, Estación Experimental de Aula Dei, Consejo Superior de Investigaciones Científicas (CSIC), Apartado 13034, 50080 Zaragoza, Spain; [2]Group of Biochemistry, Biophysics and Computational Biology (BIFI-Unizar) Joint Unit to CSIC, Edificio I+D Campus Río Ebro, 50018 Zaragoza, Spain; [3]Wheat, Sorghum, and Forage Research Unit, USDA-ARS, East Campus, University of Nebraska-Lincoln, Lincoln, NE 86583, USA; [4]Escuela Politécnica Superior de Huesca, Universidad de Zaragoza, 22071 Huesca, Spain; [5]Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, IA 50011, USA

Contents

New Phytologist (2020) **227:** 1618–1635
doi: 10.1111/nph.16444

Key words: leghemoglobins, nitrate reductase, nitric oxide, nitric oxide dioxygenase, nitrogen fixation, oxygen transport, phytoglobins, symbiosis.

Summary

Globins (Glbs) are widely distributed in archaea, bacteria and eukaryotes. They can be classified into proteins with 2/2 or 3/3 α-helical folding around the heme cavity. Both types of Glbs occur in green algae, bryophytes and vascular plants. The Glbs of angiosperms have been more intensively studied, and several protein structures have been solved. They can be hexacoordinate or pentacoordinate, depending on whether a histidine is coordinating or not at the sixth position of the iron atom. The 3/3 Glbs of class 1 and the 2/2 Glbs (also called class 3 in plants) are present in all angiosperms, whereas the 3/3 Glbs of class 2 have been only found in early angiosperms and eudicots. The three Glb classes are expected to play different roles. Class 1 Glbs are involved in hypoxia responses and modulate NO concentration, which may explain their roles in plant morphogenesis, hormone signaling, cell fate determination, nutrient deficiency, nitrogen metabolism and plant–microorganism symbioses. Symbiotic Glbs derive from class 1 or class 2 Glbs and transport O_2 in nodules. The physiological roles of class 2 and class 3 Glbs are poorly defined but could involve O_2 and NO transport and/or metabolism, respectively. More research is warranted on these intriguing proteins to determine their non-redundant functions.

I. Introduction

Hemoglobins and related heme proteins, jointly known as globins (Glbs), constitute a superfamily of proteins that are widely represented in the archaea, bacteria and eukaryotes (Vinogradov

Dedication: In memory of Bob Klucas, always eager to learn and teach about symbiotic and nonsymbiotic hemoglobins.

et al., 2005). They typically comprise a heme prosthetic group and a polypeptide of 6 to 8 α-helices named A to H. The heme is an iron-protoporphyrin IX with the ability to bind, among other ligands, diatomic gases of biological relevance such as O_2, CO and nitric oxide (NO). Binding of O_2 and CO occurs exclusively when the heme iron is in ferrous form, whereas NO can bind ferrous iron with high affinity and ferric iron with low affinity. Glbs can be classified by the spatial disposition of the α-helices relative to the heme, exhibiting 3/3-fold ('canonical') or 2/2-fold ('truncated') tertiary structures (Wittenberg *et al.*, 2002; Hoy & Hargrove, 2008). In the 3/3-helical sandwich structure, helices A, E and F are situated on one side of the heme plane and helices B, G and H lie on the other side (Fig. 1a). The 3/3 Glbs are present in metazoans, plants and many microbial eukaryotes, such as ciliates and oomycetes. The 2/2-fold is characterized by a disposition of helices B and E on one side of the heme and helices G and H on the other (Fig. 1a). The 2/2 Glbs occur in archaea, bacteria, some protists, algae and plants. Glbs can be structurally classified according to the axial coordination of the heme iron (Fig. 1b). In pentacoordinate Glbs, a His residue is at the 5th position (proximal) and exogenous ligands or water at the 6th position (distal). In hexacoordinate Glbs, a second His or, more rarely, another endogenous residue such as Lys, Gln or Tyr occupies the 6th position. Myoglobin and hemoglobin (α- and β-Glbs) of animals and symbiotic hemoglobins of legumes are pentacoordinate, whereas cytoglobin and neuroglobin of vertebrates and most Glbs of plants are hexacoordinate.

The advent of massive genome and transcriptome sequencing programmes has boosted comprehensive studies of Glb evolution (Vinogradov *et al.*, 2005; Vázquez-Limón *et al.*, 2012b; Vinogradov *et al.*, 2013; Bustamante *et al.*, 2016). Based on such studies, the Glb superfamily has been categorized into three lineages or families. The M family ('myoglobin-like', 3/3-fold) comprises two subfamilies: single-domain Glbs and flavohemoglobins. The S family ('sensor', 3/3-fold) has three subfamilies: chimeric Glb-coupled sensors, single-domain protoglobins and sensor single-domain Glbs. Finally, the T family ('truncated', 2/2-fold) can also be separated phylogenetically and structurally into three subfamilies: TrHb1, TrHb2 and TrHb3, also known as TrHbN, TrHbO and TrHbP, respectively (Wittenberg *et al.*, 2002; Vinogradov *et al.*, 2013; Bustamante *et al.*, 2016). In this context, the scope of our review will be the Glbs of plants and of their symbiotic bacterial partners. Plant Glbs are either single domain (Glbs of primitive plants and classes 1 and 2 of angiosperms) of the M family or TrHb2 (class 3) of the T family, whereas symbiotic bacteria may contain Glbs of the three families. We will describe what is known about all those classes of Glbs and will identify important challenges for future research of their structures and functions. First, the diversity of bacterial, fungal and animal Glbs are briefly outlined, providing readers with some excellent references for detailed information.

In the most recent census, 140 archaeal and 2275 bacterial genomes were examined and Glbs were detected in 23% and 52% of them, respectively (Vinogradov *et al.*, 2013). Glbs are therefore amply distributed in both kingdoms of life but are not always present. Archaea express only Glbs of the S family and TrHb1 subfamily, whereas bacteria express Glbs of all eight subfamilies. For further information see Section V, where we

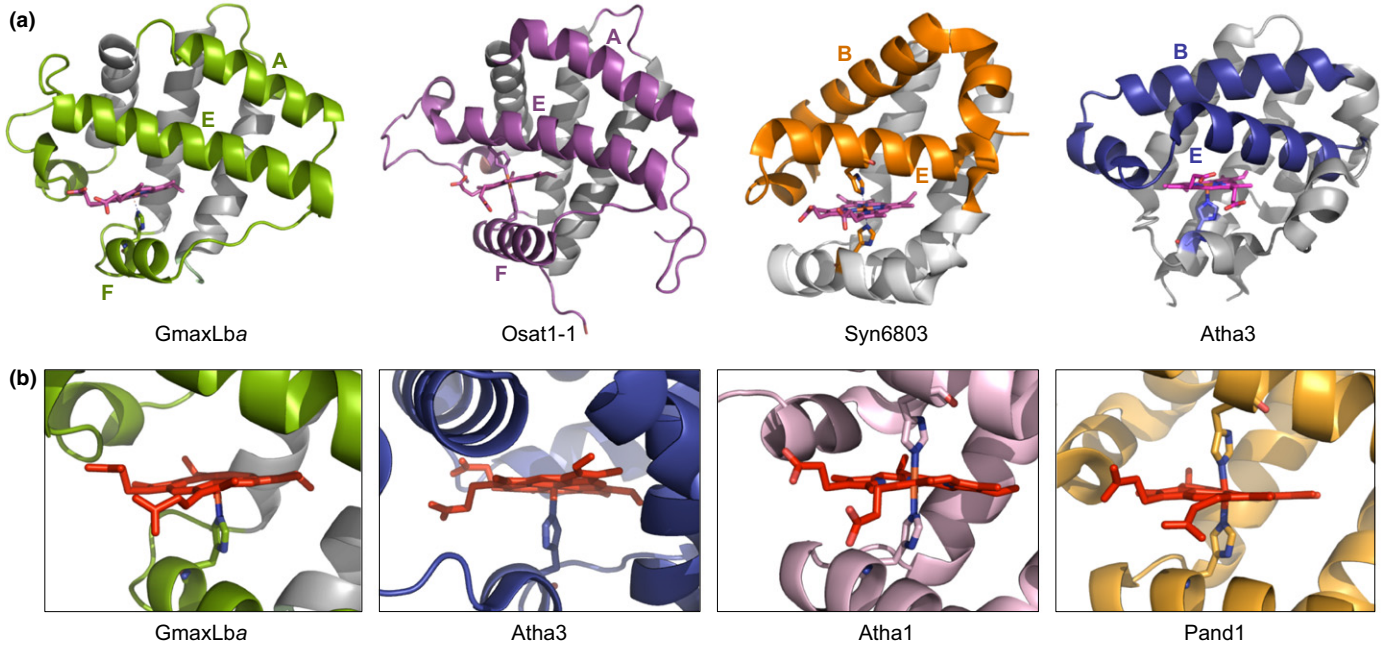

Fig. 1 Crystal X-ray structures of some globins (Glbs), illustrating (a) the 3/3-fold and 2/2-fold α-helical sandwich structures and (b) pentacoordination and hexacoordination of the heme. Abbreviations and Protein Data Bank entries for Glb sequences are given in italics. (a) Soybean leghemoglobin *a* (GmaxLb*a*; *1BIN*) and *Oryza sativa* Glb1 (Osat1-1; *1D8U*) show 3/3-fold structures. *Synechocystis* PCC 6803 (Syn6803; *1RTX*) and Arabidopsis Glb3 (Atha3; *4C0N*) show 2/2-fold structures. For simplicity, in (a) the labels of the α-helices at the bottom of the structure are omitted, and only one subunit of Osat1-1 and Atha3 is shown. (b) GmaxLb*a* and Atha3 show pentacoordinate heme, whereas Osat1-1 and *Parasponia andersonii* Glb1 (Pand1) show hexacoordinate heme.

describe the Glbs of symbiotic bacteria. In the same census, 238 genomes of fungi have been examined, 85% of which bear *Glb* genes (Vinogradov *et al.*, 2013). Notably, only genes encoding flavohemoglobins and sensor single-domain Glbs were detected. Perhaps the most studied fungal Glb is the flavohemoglobin of baker's yeast, *Saccharomyces cerevisiae* (YHB1), which is composed of an N-terminal Glb domain and a C-terminal flavin domain; the protein carries out the sequential electron transfer from NADPH to FAD to the heme and is *c.* 40% identical in both domains with the flavohemoglobin (HMP) of *Escherichia coli* (Zhu & Riggs, 1992). The widely accepted function of both flavoproteins is to prevent NO accumulation and the ensuing nitrosative stress under both aerobic and anaerobic conditions (Liu *et al.*, 2000; Gardner, 2012).

As for animal Glbs, vertebrates contain at least eight Glbs showing distinct molecular properties, preferential tissue localization and, sometimes, expression in specific taxa. Five of them also occur in invertebrates: myoglobin in muscles, hemoglobin in erythrocytes, neuroglobin in neurons, cytoglobin in fibroblasts and other cell types, and androglobin in testis cells. Three other Glbs of vertebrates have enigmatic structures and functions: GlbE in the eyes of birds, coelacanth and turtles; GlbX in fishes, amphibians and reptiles; and GlbY in amphibians (Burmester & Hankeln, 2014). Invertebrates display high heterogeneity of Glbs in primary and quaternary structures. In addition to the five Glbs shared with vertebrates, certain annelids, nematodes, crustaceans and molluscs have large multisubunit and multidomain Glbs that are frequently extracellular (Weber & Vinogradov, 2001). Notably, the complete genome sequencing of the nematode *Caenorhabditis elegans*, a classical model for developmental studies, has allowed the *in silico* identification of 33 Glbs. Mutant analysis has recently shown that one of them, Glb12, is involved in redox signaling (De Henau *et al.*, 2015). The known or surmised functions of bacterial and animal Glbs include the transport, delivery and sensing of O_2, NO metabolism, protection against nitro-oxidative stress, and enzymatic activities (Burmester & Hankeln, 2014; De Henau *et al.*, 2015). With some exceptions, however, these functions still need to be demonstrated *in vivo*.

II. Globins of cyanobacteria and algae

Cyanobacteria are an ancestral lineage of prokaryotes with the remarkable ability to perform oxygenic photosynthesis and, in some cases, nitrogen fixation. These microorganisms gave rise to plastids after an endosymbiotic event with a non-photosynthetic eukaryote that occurred *c.* 1500 million years ago (Ma) and are able to create permanent endosymbioses with ferns, bryophytes, cycads and angiosperms (Bustos-Díaz *et al.*, 2019). More than 50 genomes of cyanobacteria are now available and half of them contain *Glb* genes; the majority of the encoded products belong to the TrHb1 subfamily and a few of them to the (sensor) single-domain Glb subfamilies (Vinogradov *et al.*, 2013). The discovery more than two decades ago of a functional *Glb* gene and protein in *Nostoc commune* was a breakthrough (Potts *et al.*, 1992). The *glbN* gene is located inside the nitrogen fixation gene cluster and is expressed only during prolonged anaerobiosis under nitrogen starvation. The GlbN protein ('cyanoglobin') is shorter ('truncated') than vertebrate Glbs but similar in length to protozoan Glbs. The GlbNs of two non-nitrogen fixing cyanobacteria, *Synechocystis* PCC 6803 and *Synechococcus* PCC 7002, are also truncated and hexacoordinate, having His residues as proximal and distal heme ligands (Hoy *et al.*, 2007a; Johnson & Lecomte, 2013). In addition, both GlbNs display a rare feature: they have a non-axial His covalently bound to a heme vinyl group, which prevents heme dissociation in both ferric and ferrous forms (Hoy *et al.*, 2007b; Scott *et al.*, 2011). Cyanobacterial Glbs exhibit several types of enzyme activity *in vitro*. In the oxyferrous form ($Glb^{2+}O_2$), they can bind NO to generate NO_3^- and the ferric form (Glb^{3+}) in the NO dioxygenase (NOD) reaction (Fig. 2). In the deoxyferrous form (Glb^{2+}), cyanobacterial and plant Glbs can reduce NO_2^- to NO and hydroxylamine (NH_2OH, an intermediate of NO_2^- reduction) to NH_4^+ (Sturms *et al.*, 2011a,b). As will be described further below (Section IV.2), several enzymatic reactions, such as NOD, nitrite reductase (NiR) and hydroxylamine reductase, may be commonplace in Glbs (Fig. 2), raising intriguing questions as to their *in vivo* functions.

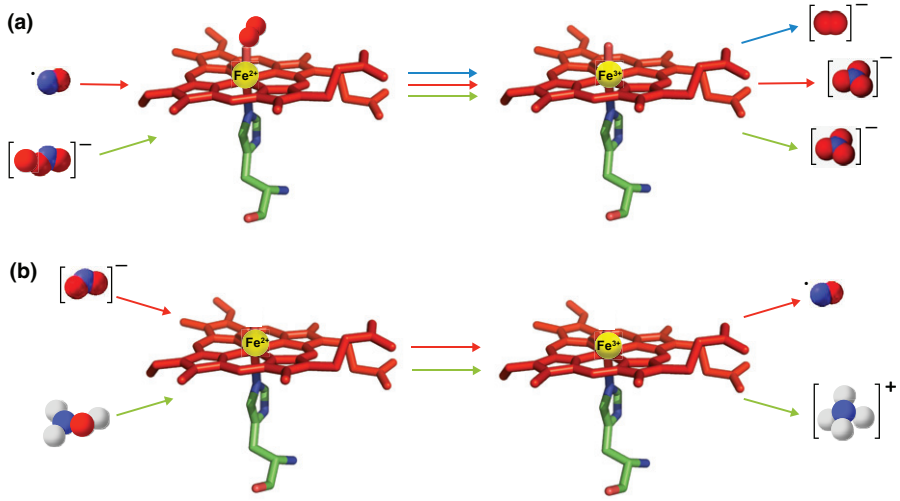

Fig. 2 Some physiologically relevant reactions of globins (Glbs) and leghemoglobins (Lbs). For simplicity, a pentacoordinate protein is shown. (a) Reactions of $Glb^{2+}O_2$ or $Lb^{2+}O_2$. The upper pathway shows autoxidation to Glb^{3+} or Lb^{3+} with production of superoxide anion radical. The middle pathway shows nitric oxide (NO) dioxygenase. The bottom pathway shows peroxynitrite ($ONOO^-$) isomerization to NO_3^-. (b) Reactions of Glb^{2+} or Lb^{2+}. The top pathway shows nitrite reductase activity producing NO. The bottom pathway shows hydroxylamine reductase activity producing NH_4^+.

The Glbs of green algae (chlorophytes) belong to the TrHb1, TrHb2 and single-domain subfamilies according to a comprehensive genome mining study of microbial eukaryotes (Johnson & Lecomte, 2013; Vinogradov *et al.*, 2013). The discovery and characterization of algal Glbs shortly followed those of cyanoglobin (Johnson & Lecomte, 2013). Two light-inducible TrHb1 genes, *LI410* and *LI637*, were identified in the unicellular green alga *Chlamydomonas eugametos*, and the LI637 protein was immunolocalized in the thylakoid membranes and the pyrenoid (Couture *et al.*, 1994). The function of this Glb was linked to regulation of photosynthesis and/or to mitigation of O_2 damage (Couture & Guertin, 1996). Most functional studies on proteins of green algae have nevertheless been conducted with the related species *Chlamydomonas reinhardtii*. For this model organism, the genomes of nuclei, chloroplasts and mitochondria have been fully sequenced and there is a large collection of mutants available (Grossman *et al.*, 2010 and references therein). The nuclear genome of *C. reinhardtii* harbors 12 genes encoding Glbs of the TrHb1 and TrHb2 subfamilies (Hemschemeier *et al.*, 2013; Phytozome v12.1). Of those, the functions of *THB1* (Cre14.g615400), *THB2* (Cre14.g615350) and *THB8* (Cre16.g661200) have been addressed with the use of mutants. The expression of *THB8* increased 1690-fold after 6 h of anoxia, and a strain in which the gene was silenced with an artificial microRNA showed impaired growth under hypoxia and increased sensitivity to NO (Hemschemeier *et al.*, 2013). These authors concluded that THB8 is part of an NO-dependent signaling pathway.

On the other hand, the expression of *THB1* and *THB2* is regulated by the nitrogen source through the participation of NIT2, a transcription factor that specifically activates NO_3^- assimilation (Johnson *et al.*, 2014; Sanz-Luque *et al.*, 2015). *THB1* is induced by NO_3^- and NO, whereas *THB2* is induced by NO_3^- but down-regulated by NO, suggesting different roles of the two genes in nitrogen metabolism (Sanz-Luque *et al.*, 2015). Detailed spectroscopic studies have unequivocally identified Lys53 as the distal heme ligand of THB1, but the biological relevance of this unusual feature in Glbs remains uncertain (Johnson *et al.*, 2014; Preimesberger *et al.*, 2017). Notably, a linkage between THB1 and nitrate reductase (NR), a molybdoenzyme crucial for NO_3^- assimilation in plants, has been established. The diaphorase activity of NR is able to efficiently reduce THB1 from its ferric to ferrous form. This reaction competes for electrons with NO_3^- reduction, thereby inhibiting (and perhaps regulating) NR activity; conversely, down-regulation of THB1 results in greater NR activity *in vivo* (Sanz-Luque *et al.*, 2015).

The diaphorase activity of NR can also donate electrons to another molybdoenzyme, the NO-forming nitrite reductase (NOFNiR), which allows the dual NR-NOFNiR system to reduce NO_3^- to NO in *C. reinhardtii* cells (Chamizo-Ampudia *et al.*, 2017). These authors have proposed an 'NO cycle' in which the NO formed from NO_3^- by NR-NOFNiR can be converted back into NO_3^- by the NOD activity of THB1. In the 'NO cycle' NR would play a central role in modulating NO concentrations (Chamizo-Ampudia *et al.*, 2017). Further support for a linkage between certain TrHbs and NR comes from another group of algae. In two species of raphidophytes, microalgae that inhabit coastal estuarine environments, a TrHb1 domain is inserted in between the cytochrome *b5* domain and the FAD domain of NR (Stewart & Coyne, 2011). This arrangement of sequences coding for two enzyme activities in a chimeric gene (*NR2-2/2HbN*) may facilitate coupling of NOD activity to the NO_2^--producing (assimilatory) activity of NR. The genomes of certain green algae and red algae (rhodophytes) encode single-domain Glbs (Vinogradov *et al.*, 2013). The picoplankton green algae *Micromonas commoda* and *Ostreococcus tauri* contain a 3/3 Glb with some features similar to land plant Glbs, such as a conserved Phe B10 and proximal and distal His residues situated at identical positions (Fernández *et al.*, 2010). To our knowledge, however, no functional studies have been carried out for any of these proteins.

III. Globins of bryophytes, lycophytes and gymnosperms

Bryophytes are non-vascular plants that colonized land *c.* 475 Ma and hence their study provides insight into the early evolution of plant Glbs. The model bryophytes for molecular research are the mosses *Physcomitrella patens* and *Sphagnum fallax* and the liverwort *Marchantia polymorpha* (Fig. 3). A search for protein sequences in the Phytozome v.12.1 database using Blastp and the class 1 Glb of Arabidopsis (Atha1) as a query reveals the presence of single-domain (3/3) Glbs in the three species. The identified Glbs were Pp3c22_15040 and Pp3c26_2330 for *P. patens*, Sphfalx0289s0010 for *S. fallax*, and Mapoly0014s0083, Mapoly0104s0016 and Mapoly0104s0018 for *M. polymorpha*. One 3/3 Glb sequence of the moss *Ceratodon purpureus* is also available (GenBank accession no. AAG22831). The *Glb* genes of bryophytes contain three introns at identical positions to those of vascular plants, suggesting that the ancestral *Glb* gene of plants was also interrupted by three introns (Garrocho-Villegas & Arredondo-Peter, 2008). These authors showed that the *Glb* gene of *C. purpureus* is functional and expresses in gametophytes and protonemas under various physiological and stressful conditions. The Glbs of *P. patens* and *C. purpureus* have been purified in recombinant form and their UV-visible spectra revealed that they are mostly hexacoordinate and can bind O_2 (Vázquez-Limón *et al.*, 2012a). Some 2/2 Glbs of the TrHb2 subfamily were also found in the Phytozome v12.1 database using BLASTP and the class 3 Glb of Arabidopsis (Atha3) as a query. The identified Glbs were Pp3c1_3380 and Pp3c6_250 for *P. patens* and Sphfalx0029s0049 and Sphfalx0052s0112 for *S. fallax*. However, none was observed for *M. polymorpha* in the same database or for *C. purpureus* in GenBank. The information about the Glbs from pteridophytes and gymnosperms is even more scarce. The model species are *Selaginella moellendorffii* and *Picea sitchensis*, respectively. Both have 3/3 and 2/2 Glb genes (see accession nos. in the legend of Fig. 3), but the structures and functions of the predicted proteins are a mystery.

IV. Globins of angiosperms

The Glbs of angiosperms have been by far the most extensively studied. The first plant Glb was identified in 1939 by Kubo in soybean (*Glycine max*) nodules and later named leghemoglobin

3/3 globins

	Mcom	Ppat	Mpol	Smoe	Psit	Bdis1-1	Bdis1-2	Atha1	Atha2
Ppat	28 49								
Mpol	36 49	50 66							
Smoe	34 50	47 65	48 67						
Psit	32 53	47 66	47 65	51 66					
Bdis1-1	37 55	50 67	47 63	51 64	61 75				
Bdis1-2	32 47	47 68	46 64	47 63	58 73	66 78			
Atha1	29 52	49 69	48 62	50 67	61 78	70 84	62 74		
Atha2	34 56	44 66	47 64	48 64	51 72	56 70	53 70	55 71	
LjapLb1	33 50	39 57	39 52	38 55	43 60	42 63	45 61	42 57	49 71

2/2 globins

	CreiTHB2	Ppat3-1	Smoe3	Psit3	Bdis3	Atha3	Ljap3-1
Ppat3-1	20 37						
Smoe3	30 51	70 81					
Psit3	27 45	68 77	77 87				
Bdis3	27 40	58 72	65 80	69 82			
Atha3	23 42	63 78	73 86	74 85	70 80		
Ljap3-1	NS NS	61 76	65 80	70 83	70 81	74 82	
Ljap3-2	NS NS	64 80	71 82	74 87	73 85	78 85	83 90

Fig. 3 Sequence homologies of 3/3 globins and 2/2 globins from model algae, bryophytes and vascular plants. Numbers in blue and red are percentages of identity and similarity, respectively. These were calculated with BLASTP of the National Center for Biotechnology Information. Abbreviations and accession nos. are as follows. Arabidopsis (Atha1, At2g16060; Atha2, At3g10520; Atha3, At4g32690); *Brachypodium distachyon* (Bdis1-1, Bradi1g69320; Bdis1-2, Bradi2g19690; Bdis3, Bradi1g37100); *Chlamydomonas reinhardtii* (CreiTHB2, Cre14.g615350); *Lotus japonicus* (LjapLb1, Lj5g3v0035290.2; Ljap3-1, Lj1g3v2035270; Ljap3-2, Lj1g3v0948590); *Marchantia polymorpha* (Mpol, Mapoly0104s0016); *Micromonas commoda* (Mcom, XP_002504250); *Physcomitrella patens* (Ppat, XP_024360203.1; Ppat3-1, XP_024377531); *Picea sitchensis* (Psit, ABR17163; Psit3, ABK22150); *Selaginella moellendorffii* (Smoe, EFJ10590; Smoe3, XP_002991488); NS, no significant homology.

(Lb) from 'legume hemoglobin' (for historical details see reviews by Appleby, 1984 and Garrocho-Villegas *et al.*, 2007). In the following years Lbs were found in other legume nodules and symbiotic Glbs in nodules of actinorhizal plants (Section V). By the late 1980s it was evident that all vascular plants contain a second type of Glbs, termed 'nonsymbiotic' because of their widespread expression in plant organs and species (Landsman *et al.*, 1986). Afterwards, 'nonsymbiotic hemoglobins' were categorized into two phylogenetic classes. The class 1 Glbs of barley (*Hordeum vulgare*, Hvul1-1; Duff *et al.*, 1997; Das *et al.*, 1999), rice (*Oryza sativa*, Osat1-1; Arredondo-Peter *et al.*, 1997), soybean (Gmax1; Andersson *et al.*, 1996) and Arabidopsis (Atha1; Trevaskis *et al.*, 1997) were purified in recombinant form, characterized and, in certain cases, crystallized (Section IV.1). A class 2 Glb of Arabidopsis (Atha2) was simultaneously identified (Trevaskis *et al.*, 1997). A few years later, a third class of Glb (Atha3) was discovered and found to have a 2/2-fold structure and high homology to the truncated Glbs of bacteria (Watts *et al.*, 2001). Class 3 Glbs belong to the TrHb2 subfamily and are also known as 'truncated' despite the fact that they are usually longer than the other two classes of Glbs. Recently, the term 'phytoglobin' was recommended for the three classes of plant Glbs, keeping 'leghemoglobin' for historical reasons (Hill *et al.*, 2016). For simplicity and consistency with other globins, in this review we will also use the abbreviation 'Glb' to refer to phytoglobins.

1. Structure of globins

The presence of three Glb classes in angiosperms suggests distinct functions. Dicots contain members of each of them, whereas monocots have only class 1 and class 3 Glbs. A great deal of experimental work has been directed at defining the structure and function of Lbs and of class 1 and 2 Glbs, whereas much less attention has been paid to class 3 Glbs. Atomic resolution 3D structures have been measured for several class 1 Glbs and Lbs and a single class 3 Glb. The structures (Protein Data Bank entries in italics) have been solved for Osat1-1 of rice (*1D8U*; Hargrove *et al.*, 2000), Zmay1-1 of maize (*Zea mays*; *2R50*; Smagghe *et al.*, 2009), Atha1 of Arabidopsis (*3ZHW*; Mukhi *et al.*, 2013), Hvul1-1 of barley (*2OIF*; Hoy *et al.*, 2007a), Ttom1 of *Trema tomentosa* (*3QQQ*; Kakar *et al.*, 2011) and Pand1 of *Parasponia andersonii* (*3QQR*; Kakar *et al.*, 2011). These are all 3/3 structures with reversible hexacoordination of the heme iron by two endogenous His side chains (Fig. 1). There are currently no structures of class 2 Glbs, but their sequence similarities to class 1 Glbs predict that they would be comparable in the reversibility of the hexacoordinated distal His. The structure of Atha3 has been solved more recently (*4C0N*; Reeder & Hough, 2014; Mukhi *et al.*, 2016). Atha3 exhibits the expected 2/2 fold and is dimeric and pentacoordinate (Fig. 1), with a large distal heme pocket lacking any side chains that could coordinate the heme iron. There are no His side chains in the distal pocket, but there are Tyr and Trp side chains *c.* 3 Å from the ligand binding site that could interact with a bound ligand.

Structural analysis of symbiotic Glbs reveals that those originating from class 2 Glbs such as Lbs and Cgla2 have evolved pentacoordinate heme pockets similar to O_2-transport Glbs in animals; however, the control of O_2 binding is dictated by proximal coordination rather than by direct interactions between the distal His and bound O_2 (Smagghe *et al.*, 2009). Unlike Lbs, Pand1 derives from a class 1 Glb and is hexacoordinate as Glb^{3+} and pentacoordinate as Glb^{2+}; like Lbs, however, it has an O_2 dissociation rate constant commensurate with its function as an O_2 transporter (Kakar *et al.*, 2011). The O_2 dissociation rate constants of O_2 transporters are generally $> 5\ s^{-1}$, while most Glbs that are not O_2 transporters have values $< 1\ s^{-1}$. The value for Pand1 is $15\ s^{-1}$, while Ttom1 (a close nonsymbiotic relative with

93% amino acid identity) is 0.38 s^{-1} (Sturms *et al.*, 2010). These data reiterate the importance of rapid O_2 dissociation for O_2 transport and suggest that ferric hexacoordination and a lower O_2 dissociation rate constant might be important for Glb function.

2. Reactivity of globins

The general reactivity of Glbs in reversible ligand binding and redox chemistry is both a blessing and a curse in terms of the discovery of biological function. The blessing of general reactivity is the detailed biochemistry that can result from experiments *in vitro*; the curse is the necessity of verifying that such reactivity is biologically relevant before offering it in support of a hypothesis for function.

Reversible distal histidine coordination Reversible distal His coordination was first recognized as a natural phenomenon in Hvul1-1 and Osat1-1 (Arredondo-Peter *et al.*, 1997; Duff *et al.*, 1997). The affinity constant for distal His binding is larger in Glb^{3+} than in Glb^{2+} and also larger in class 2 compared to class 1 Glbs (Smagghe *et al.*, 2009). Prior to this discovery, the *bis*-histidyl heme structure was considered relatively inert toward exogenous ligand binding and useful mainly for electron transfer proteins such as cytochrome b_5. Hexacoordinate Glbs are also widespread in animals and in some bacteria (Kakar *et al.*, 2010). In general, the affinity constants for His binding are larger in the hexacoordinate Glbs of animals than in those of plants.

There are three important consequences of hexacoordination. First, it lowers the midpoint redox potential for the protein compared to pentacoordinate Glbs because the His binds more tightly to Glb^{3+} than to Glb^{2+}, thus favoring the oxidized form of the protein. This has been observed in generally lower midpoint redox potentials for many of the plant Glbs (Halder *et al.*, 2007; Mot *et al.*, 2018). Second, it increases rate constants for electron transfer to and from the heme (Weiland *et al.*, 2004) because there is minimal structural change in switching between oxidation states. These first two factors are relevant to mechanisms for protein reduction and electron transfer to bound ligands, and certainly contribute to the generally faster rate constants for autooxidation (Weiland *et al.*, 2004). Finally, it lowers affinity constants for other ligands by competing for the active site. In class 2 Glbs, where affinity for the distal His is large, the affinity for O_2 and CO are accordingly much lower. This provides a potential mechanism for regulating ligand affinity through the strength of distal His coordination (Smagghe *et al.*, 2009).

Reversible ligand binding Dissociation of the bound His allows access to the distal heme pocket that can bind reversibly to many small ligands. Accordingly, Glb^{2+} will reversibly bind O_2, CO and NO. CO binding kinetics have been used to measure the rate and equilibrium constants for reversible distal His coordination and to displace O_2 for measurement of O_2 dissociation rate constants. Such studies have shown that the *c.* 50-fold higher equilibrium affinity constant for the distal His in class 2 Glbs is due largely to the increased association rate constant. Reversible O_2 binding must compete with distal His coordination, and thus the larger internal competition for His binding in class 2 Glbs lowers its affinity compared to class 1 Glbs. Another effect of reversible His coordination is a kinetic limit to exogenous ligand binding to the heme iron. If fractional saturation of hexacoordination is *c.* 1, the exogenous ligand must wait for the bound His to dissociate before it can bind (Smagghe *et al.*, 2009). The O_2 dissociation rate constants in class 1 and class 2 Glbs are slow, averaging 0.14 and 1.1 s^{-1}, respectively, compared to known O_2 transporters (Smagghe *et al.*, 2009). Comparatively little is known about reversible ligand binding in class 3 Glbs (Mukhi *et al.*, 2016). The O_2 dissociation rate constant for Atha3 (0.35 s^{-1}) is in between those of Atha1 and Atha2, but the association rate constants for both CO and O_2 are very low in comparison. This is surprising in light of the large open binding pocket, observed in the structure, suggesting that ligand access to the pocket could be limiting. Regulation of ligand binding in support of function has been proposed by the detailed study of bimolecular and geminate ligand rebinding to Atha1 and Atha2 (Bruno *et al.*, 2007; Spyrakis *et al.*, 2011). Recently, extensive kinetic studies on the various Glbs of the model legume *Lotus japonicus* (Calvo-Begueria *et al.*, 2017) and Arabidopsis (Mot *et al.*, 2018) strongly suggest non-redundant functions.

Nitric oxide dioxygenase activity This is the core biochemical reaction grounding most of the functional hypotheses of Glbs described below (Section IV.3). The reaction takes place between the oxyferrous form ($Glb^{2+}O_2$) and NO to yield NO_3^- (Fig. 2a). Its biological relevance was initially recognized in bacterial flavohemoglobins as a mechanism to detoxify NO (Gardner *et al.*, 1998). The heme domain binds O_2 and reacts with NO, and the reductase domain shuttles electrons sequentially to the heme to continually reduce it back to the ferrous state. A thoughtful and thorough review of NOD and Glbs in light of this prevalent hypothesis for single domain Glbs has been published previously (Gardner, 2012).

Other redox reactions of Glbs In an effort to explore anaerobic biochemistry in Glbs, it was discovered that class 1 Glbs and *Synechocystis* TrHb reduce NO_2^- to NO (NiR activity) at a rate that is > 10-fold faster than other (mammalian) Glbs (Sturms *et al.*, 2011b; Tiso *et al.*, 2012). The product of the reaction *in vitro* is $Glb^{2+}NO$ and Glb^{3+}, but small amounts of gas-phase NO were also detected, leaving open the possibility that NO_2^- reduction by Glbs could generate NO under the right conditions (Tiso *et al.*, 2012). The rate constants for the reactions of the three Glbs of sugar beet (*Beta vulgaris*) with NO and NO_2^- have been found to be distinctly different, hinting at non-redundant roles in nitrogen metabolism (Leiva-Eriksson *et al.*, 2019). Other reactions of Glbs are the peroxidase activity (Sakamoto *et al.*, 2004; Violante-Mota *et al.*, 2010) and the anaerobic reduction of NO_2^- and NH_2OH (Sturms *et al.*, 2011a,b). Peroxidase reactions, like reversible ligand binding and the NO-oxidizing half of the NOD cycle, are a common behavior of Glbs that have been oxidized by peroxides. In fact, these reactions are even called 'pseudoperoxidase' reactions in light of the fact that they are not given much credence *in vivo* (Alayash, 2001). *In vitro* peroxidase rates for Lbs and Glbs are measurable but considered unlikely to be physiologically relevant.

Nitrate is a respiratory electron acceptor in many bacteria. Plants lack a respiratory NR yet benefit from NO_3^- during hypoxia (Oliveira *et al.*, 2013). Glbs do not react with NO_3^- and their reactions with NO_2^- are described at the start of the previous paragraph. The next stable nitrogen metabolite in the ammonification pathway is NH_2OH. The reduction of NH_2OH to NH_4^+ involves the transfer of two electrons (Fig. 2b) and thus requires intermolecular electron exchange between Glb subunits (Athwal *et al.*, 2015). This reaction could potentially serve anaerobic metabolism by providing a step along the pathway of fermentative ammonification. Reactions between class 1 Glbs and NH_2OH were tested and found to be two to three orders of magnitude faster than those with animal Glbs (Sturms *et al.*, 2011a). Unlike NO_2^- reduction, the reaction with NH_2OH can be sustained catalytically at a rate limited only by the dissociation of the distal His from the ferrous heme iron. Further, it requires His hexacoordination for rapid reduction (Athwal *et al.*, 2016; Alagurajan *et al.*, 2018).

3. Functional analysis of globins

The first hypothesis for Glb function was O_2 transport, based solely on the intuitive association of animal hemoglobin and myoglobin with O_2 transport. However, Glbs were shown to lack the appropriate concentrations and kinetic and affinity constants, and therefore the search for a function quickly moved away from O_2 transport. It was then proposed that Glbs could sense a decrease in O_2 concentration in the roots (Appleby *et al.*, 1988). Although in general Glbs have very low O_2 dissociation ('off') rates, some of them could function as O_2 transporters or sensors under certain growth conditions or in specific tissues (Arredondo-Peter *et al.*, 1998). Subsequent studies with maize cultured cells overexpressing Hvul1-1 concluded that class 1 Glbs may maintain the energy status of the cells under hypoxic conditions by promoting the glycolytic flux through NADH oxidation (Sowa *et al.*, 1998). This beneficial effect would be in keeping with a subsequent study showing that Arabidopsis plants overexpressing Atha1 or expressing Pand1 show enhanced survival under hypoxia and more vigorous early growth (Hunt *et al.*, 2002).

The NOD reaction was put forward to explain some Glb functions. Perazzolli *et al.* (2004) showed that the oxyferrous form of Atha1 reacts with NO and, even though this was expected for any Glb, they interpreted the result as support for the NOD hypothesis. In this scenario the NO generated during hypoxia is converted back to NO_3^- via NOD by Glb, aiding fermentation by oxidizing NAD(P)H in the process (Igamberdiev & Hill, 2004). This would explain the benefit of NO_3^- to plants during hypoxia in providing a mechanism for generating ATP under anaerobic conditions. It would also serve naturally hypoxic environments such as seeds (Matilla & Rodríguez-Gacio, 2013). It does not explain why limited O_2 during hypoxia would be prioritized for usage by NOD instead of respiration, but the complexities of hypoxic NO metabolism certainly leave room for this result.

Physiological work using transgenic plants with altered concentrations of Glbs has shown that these proteins play an important role in plant development and morphogenesis, and largely supports the theory that they do so by modulating NO concentrations (Hill, 2012; Hebelstrup *et al.*, 2013). Glbs may act through NO modulation to affect the signaling pathways of hormones including auxins (Elhiti *et al.*, 2013), cytokinins (Ross *et al.*, 2004; Wang *et al.*, 2011), salicylate (Mur *et al.*, 2012), jasmonates (Mur *et al.*, 2012; Mira *et al.*, 2016), ethylene (Mur *et al.*, 2012; Kapoor *et al.*, 2018; Hartman *et al.*, 2019) and abscisic acid (ABA; Kapoor *et al.*, 2018; Rubio *et al.*, 2019). Two exemplary studies, concerning the response of Arabidopsis to biotic and abiotic stress, will illustrate the triple crosstalk between Glbs, hormones and NO. The first study concluded that plants deficient in Atha1, but not in Atha2 or Atha3, were more resistant to attack by *Pseudomonas syringae*, whereas the *Atha1* overexpressing line had a compromised defense response; these effects were mediated by changes in the concentrations of NO, jasmonic acid and ethylene (Mur *et al.*, 2012). The second study showed that ethylene increases Atha1 which, by depleting NO, limits proteolysis of the group VII ethylene response factor (ERFVII) transcription factor under normoxic conditions; the stabilized ERFVII is then translocated to the nucleus, where it induces expression of hypoxia-related genes under O_2 limitation, which in turn increases hypoxia tolerance of the root and shoot apical meristems (Hartman *et al.*, 2019).

Strong evidence has accumulated that the cellular levels and localization of Glbs determine cell fate (Huang *et al.*, 2014; Stasolla *et al.*, 2019). These authors showed that the two maize Glbs are expressed in different cellular groups and respond distinctly during somatic embryogenesis. Suppression of *Zmay1* causes massive programmed cell death of embryonic cells leading to abortion, whereas suppression of *Zmay2* results in programmed cell death only of basal cells, thereby releasing the immature embryos and allowing them to develop further; these effects are thought to be mediated by NO, ABA and ethylene (Stasolla & Hill, 2017; Stasolla *et al.*, 2019). All these observations raise major questions. Are the promoters of each *Glb* gene the only factor that dictates the specificity of the protein activity? How are Glbs able to distinguish between the signaling pathway of each hormone? Do all Glb functions involve NO? With such a diversity of Glbs, hormones and heme ligands, and hence of potential interactions among them, it is clear that the biochemical properties of Glbs (besides their expression profiles) and their *in vivo* ligands (besides NO) will be of particular interest in future research.

There is solid support for the NO scavenging reaction of Glbs *in vivo*. In Arabidopsis the reaction can be carried out by both Atha1 and Atha2 (Hebelstrup & Jensen, 2008; Kuruthukulangarakoola *et al.*, 2017) and has been observed in stomatal cells, where a crosstalk between Glbs, NO and ABA may occur (Rubio *et al.*, 2019). Recent work with Arabidopsis and barley has shown that NO can be 'fixed' into nitrogen metabolism in plants overexpressing, respectively, *Atha1* or *Atha2* (Kuruthukulangarakoola *et al.*, 2017) and *Hvul1-1* (Zhang *et al.*, 2019). The authors suggest that this mechanism may be useful for plants in soils with poor nitrogen availability, but the relevance of these findings under physiological, nitrogen-sufficient conditions remains uncertain. Notably, Glb expression is dependent on nutrient status. In Arabidopsis under hypoxic conditions, NO_3^- nutrition increases *Atha1* mRNA level, NR activity, NO, fermentative metabolism and ATP production, whereas NH_4^+ nutrition has no effect (Wany *et al.*, 2019). The

metabolic changes with NO_3^- have been attributed to the generation of NO from NO_3^- by NR and the electron transport chain, but the distinct response of Glbs in plants fed with different sources of nitrogen deserves further investigation. Nutrient deficiencies also affect Glbs expression. In tomato (*Solanum lycopersicum*), the removal of phosphorus, potassium or iron from the hydroponic culture induces the expression of *Slyc1* but not *Slyc2* (Wang *et al.*, 2003). In rice, the overexpression of *Osat2* improves root growth under potassium deficiency and reduces the production of reactive oxygen species (Shankar *et al.*, 2018). The mechanism linking Glbs with a better performance of plants under nutrient stress is unknown, but it may involve, at least to a certain extent, the ability of Glbs to modulate the concentrations (or signaling pathways) of H_2O_2 (Yang *et al.*, 2005; Shankar *et al.*, 2018) and NO through NOD activity (Igamberdiev & Hill, 2004; Perazzolli *et al.*, 2004). However, some functions of class 2 Glbs are most probably unrelated to NO. It has been suggested that these proteins are involved in O_2 transport based on the observations that their rate constants for O_2 binding are similar to those of symbiotic Glbs (Spyrakis *et al.*, 2011). In support of this hypothesis, Atha2 participates in fatty acid metabolism and in the accumulation of polyunsaturated fatty acids by facilitating O_2 supply in developing seeds (Vigeolas *et al.*, 2011).

The most recurrent hypothesis for Glb function rests on the NOD reaction. Gardner (2012) has defined criteria for identifying biological NOD activity in a Glb: 'Full proof of a NOD function requires the demonstration of catalytic O_2-dependent NO metabolism within the native organism and a specific electron donor needs to be identified for demonstration of efficient catalysis *in vitro*'. While there is strong evidence in support of a role of Glbs in modulation of NO concentrations by the reaction of $Glb^{2+}O_2$ with NO, the search for specific electron donors of Glb^{3+} has been a bumpy road with no end. Several mechanisms for Glb^{3+} reduction have been suggested. Igamberdiev *et al.* (2006) proposed that ascorbic acid reduces Glb^{3+} in a process facilitated by monodehydroascorbate reductase (MDHAR). In this hypothesis MDHAR replenishes the ascorbic acid pool using electrons from NAD(P)H to keep functional Glb^{2+}. Subsequent detailed analysis of reactions between ascorbic acid, Glb and MDHAR refuted this hypothesis by showing that rates of Glb^{3+} reduction by ascorbic acid are too slow, and that MDHAR would likely have little effect on them under physiological concentrations of ascorbic acid (Wang & Hargrove, 2013). Other studies showed that Glb^{3+} may be reduced by free flavins acting as electron carriers from NADH, although the efficiency was high for class 1 and class 2 Glbs but relatively low for class 3 Glbs (Sainz *et al.*, 2013). Flavins are present in the nuclei, cytoplasm and plastids, where Glbs are located (Section IV.4), and hence these electron carriers appear to be suitable for reducing some classes of Glb^{3+} *in vivo*. Also, the NR of *C. reinhardtii* provides electrons to THB1 with high efficiency (Sanz-Luque *et al.*, 2015), but it remains to be seen if cytosolic NR can reduce also Glb^{3+} in vascular plants. Another flavoprotein that may reduce Glb^{3+} is ferredoxin $NADP^+$ oxidoreductase. In a yeast complementation test, the enzyme of poplar (*Populus tremula* x *tremuloides*) restores the resistance of mutant yeast cells to NO toxicity when co-expressed with poplar Glb1 but not with Glb3, suggesting a specific interaction between the reductase and Glb1 (Jokipii-Lukkari *et al.*, 2016). The *in vivo* functioning of this Glb^{3+} reductase deserves detailed scrutiny in other model species.

Recently, the term 'metphytoglobin reductase' has been coined in reference to a specific reductase activity for Glb^{3+} (Gupta & Igamberdiev, 2016) and to a general NAD(P)H oxidation activity in crude extracts in support of the NOD activity of Glbs (Wany *et al.*, 2019). Such a specific enzyme has not been identified yet and any reference using this term is therefore misleading. Thus, the search for a reductase of Glb^{3+} continues and, until identified, mechanisms for NO modulation other than NOD should be also considered.

4. Localization of globins in plants

The discovery of multiple classes and subclasses of Glbs in angiosperms brought immediate attention to the gene expression profiles in plant organs, tissues and cells. These studies are too numerous to be cited and a few examples may suffice to appreciate their complexity. The first expression analysis was conducted in barley, where *Hvul1-1* mRNA was localized in aleurone layers exposed to anaerobic conditions and in roots experiencing flooding stress (Taylor *et al.*, 1994). Detailed studies by Trevaskis *et al.* (1997) and Hunt *et al.* (2001) showed that *Atha1* is expressed in the hypocotyls and cotyledons of germinating and very young seedlings and is induced in the roots and rosette leaves by low O_2 concentrations; by contrast, *Atha2* is not inducible by hypoxia and is expressed in the bolt stem, roots, leaves and inflorescences of mature flowering plants. Two other examples of contrasting expression patterns of Glbs are *L. japonicus* and sugar beet. The genome of *L. japonicus* harbors two class 1, one class 2 and two class 3 Glb genes. *Ljap1-1*, *Ljap2* and *Ljap3-1* are highly expressed in nodules, *Ljap1-2* in leaves and *Ljap3-2* throughout the plant (Bustos-Sanmamed *et al.*, 2011). Sugar beet has two class 1 Glb genes and one class 2 Glb gene. *Bvul1-1* is highly expressed in hypocotyls and flowers, *Bvul1-2* in seeds, and *Bvul2* in hypocotyls, cotyledons and leaves, whereas neither of the three genes is significantly expressed in roots (Leiva-Eriksson *et al.*, 2014). Although the experimental conditions of the mentioned studies vary considerably, it is evident that there are not consistent expression patterns of each Glb class across plant species, or even between the two class 1 Glbs in the case of *L. japonicus* and sugar beet.

The localization of Glbs has been also been investigated at the subcellular level. An early report suggested a cytoplasmic localization and increased accumulation in differentiating tracheids (Ross *et al.*, 2001). Other studies conclude that Glbs are preferentially localized to the nuclei and cytoplasm. This is the case in alfalfa (*Medicago sativa*) suspension cells (Seregélyes *et al.*, 2000), Arabidopsis roots and leaves (Hebelstrup *et al.*, 2008), and *L. japonicus* roots, leaves and nodules (Sainz *et al.*, 2013; Rubio *et al.*, 2019). Glbs are absent from mitochondria, but there are indications that Glbs occur in chloroplasts. In wheat (*Triticum aestivum*), a putative interaction of a class 3 Glb with photosystem I and II subunits has been reported (Kim *et al.*, 2013). Also, Bvul1-1 contains a chloroplast transit peptide (Leiva-Eriksson *et al.*, 2014).

Very recent work using quantitative immunogold labeling of overexpressing and silenced/knockout lines of Arabidopsis, combined with confocal microscopy of Glbs tagged with green fluorescent protein, has shown the presence of several Glbs in the chloroplasts and amyloplasts (Rubio *et al.*, 2019). These Glbs might play a role associated with photosynthesis and/or NO homeostasis because chloroplasts are a source of NO within the cell (Jasid *et al.*, 2006).

V. Globins in symbiosis

The function of Lbs is the transport and delivery of O_2 to the bacteroids at a low but stable concentration in the cytosol of infected cells (Appleby, 1984). This avoids the irreversible inactivation of O_2-labile nitrogenase while sustaining an efficient respiration of bacteroids to produce energy for N_2 fixation. The Glbs from nodules of *P. andersonii* (Pand1; Appleby *et al.*, 1983) and of the actinorhizal plants *Casuarina glauca* (Cgla2; Fleming *et al.*, 1987) and *Myrica gale* (Mgal; Pathirana & Tjepkema, 1995) have been characterized. The non-legume *Parasponia* is unique in forming nodules with bradyrhizobia, whereas actinorhizal nodules are elicited by *Frankia*. Unlike other actinorhizal nodules, *Frankia* strains do not form vesicles to protect nitrogenase in *Casuarina* nodules, which may be linked to the abundance of Cgla2 in the cytosol (Fleming *et al.*, 1987; Silvester *et al.*, 2008). By contrast, a Glb has been found at low concentration in nodules of *Alnus glutinosa*, although it may be a truncated Glb of *Frankia* (Silvester *et al.*, 2008). Interestingly, Pand1 is similar to Lbs in many critical aspects, including the kinetics for O_2 and CO binding (Kortt *et al.*, 1988; Sturms *et al.*, 2010) and the localization in the cytosol of nodule infected cells (Trinick *et al.*, 1989). Thus, Pand1 transports O_2 to the bacteroids but may play also a nonsymbiotic role because it is expressed in the roots. By contrast, the expression of Pand2 is low in most tissues and hence its function remains elusive (van Velzen *et al.*, 2018). A dual role has been proposed also for Glb1 of *Alnus firma* (Sasakura *et al.*, 2006).

The use of RNA interference to silence the expression of the three Lbs of *L. japonicus* demonstrated that they are essential for symbiotic N_2 fixation but not for the growth of plants fed on nitrogen fertilizer (Ott *et al.*, 2005). The mutant nodules showed an increase in the internal free O_2 concentration and alterations in the expression of many genes, including bacterial *nif* and *fix* (Ott *et al.*, 2009). Subsequent work using CRISPR/Cas9 to obtain single, double and triple (*lb123*) mutants concluded that the three Lbs act additively to maintain optimal N_2 fixation (Wang *et al.*, 2019). This study also provided RNA-sequencing information of *c.* 20 000 genes of the *lb123* nodules. These Lb-free nodules showed cell vacuolization, accumulation of poly-β-hydroxybutyrate, disruption of mitochondria and enhanced concentrations of superoxide and peroxide (Wang *et al.*, 2019).

The existence of mechanisms for Lb^{3+} reduction was postulated on the basis that only the Lb^{2+} state is functional and that there are conditions in nodules, such as the slightly acid pH, that are favorable for Lb^{2+} (auto)oxidation. Free flavins, at physiological nodule concentrations, mediate the efficient NADH-dependent reduction of Lb^{3+} to $Lb^{2+}O_2$ aerobically and of Lb^{3+} to Lb^{2+} anaerobically (Becana & Klucas, 1990). Also, a flavoprotein with the required high affinity for Lb^{3+} ($K_m = 9\ \mu M$) and NADH ($K_m = 50\ \mu M$) was purified from soybean nodules and named ferric leghemoglobin reductase (Ji *et al.*, 1991). The enzyme required O_2 and generated superoxide radical (Becana & Klucas, 1990). The subcellular localization of the enzyme was not ascertained, and thus its operativity *in vivo* requires further investigation.

Soybean nodules contain *c.* 80% of Lb^{2+}, *c.* 20% of $Lb^{2+}O_2$ and tiny amounts of Lb^{3+} (Lee *et al.*, 1995). However, additional forms of Lb have been detected under certain developmental or stressful conditions. The presence of nitrosyl-Lb in soybean intact nodules was reported using electron paramagnetic resonance in healthy nodules (Mathieu *et al.*, 1998) and in nodules deficient in the bacteroid denitrification enzymes NirK or Nor (Calvo-Begueria *et al.*, 2018), especially in response to flooding conditions (Sánchez *et al.*, 2010). On the other hand, green pigments derived from Lbs were observed in senescing soybean nodules (Jun *et al.*, 1994a,b) and almost 20 years later identified as nitri-Lbs, unveiling the generation of nitrating reactive compounds in nodules (Navascués *et al.*, 2012). By contrast, Lbs bearing nitro-Tyr are more abundant in young nodules, suggesting that they originate as a result of active nodule metabolism (Sainz *et al.*, 2015). Lbs could act as scavengers of nitrogen dioxide ($^{\cdot}NO_2$) and peroxynitrite ($ONOO^-$), which cause Tyr nitration and cytotoxicity (Herold & Puppo, 2005; Sainz *et al.*, 2015). *In vitro* experiments have shown that $Lb^{2+}O_2$ reacts with NO (NOD activity) and $ONOO^-$ producing Lb^{3+} and NO_3^- (Herold & Puppo, 2005; Fig. 2a).

As occurs with other plant organs, Glbs are expressed in nodules (Uchiumi *et al.*, 2002; Vieweg *et al.*, 2005; Bustos-Sanmamed *et al.*, 2011; Berger *et al.*, 2020). Several studies have proposed a role of Ljap1-1 in decreasing NO concentrations in roots after infection by its natural symbiont *Mesorhizobium loti*, thereby preventing the plant's defense response (Nagata *et al.*, 2008; Fukudome *et al.*, 2016). Ljap1-1 is the only Glb of *L. japonicus* that is induced by NO or its immediate precursor NO_2^- (Shimoda *et al.*, 2005). Plants overexpressing *Ljap1-1* or *A. firma* Glb1 exhibit better symbiotic performance and produce nodules with lower NO concentrations than the wild-type plants (Shimoda *et al.*, 2009; Fukudome *et al.*, 2019). Similar beneficial effects of overexpressing *Mtru1-1* have been very recently reported in the *Sinorhizobium meliloti–Medicago truncatula* symbiosis (Berger *et al.*, 2020).

The bacterial symbionts also have several classes of functional Glbs: *S. meliloti* and *M. loti* have flavohemoglobins (Hmp), whereas *Bradyrhizobium japonicum* has a single-domain Glb (Bjgb). In addition, all these rhizobial species may have TrHbs. This is also the case for *Frankia* strain CcI3, which produces two truncated Glbs of different subfamilies, HbN (TrHb1) and HbO (TrHb2). HbN may be involved in protection against nitrosative stress and HbO in adaptation to low O_2 concentrations (Niemann & Tisa, 2008; Coats *et al.*, 2009). Interestingly, the HbO of *Frankia* and a class 3 Glb of its host plant *Datisca glomerata* are induced in symbiosis (Pawlowski *et al.*, 2007).

The Hmp of rhizobia can modulate NO concentrations in nodules. Elegant studies using an *S. meliloti* strain overexpressing the *hmp* gene as well as *M. truncatula* hairy roots overexpressing the *hmp* gene under the control of a nodule-specific promoter

demonstrated that NO is produced during the early stages of interaction with the bacteria and that it is required for the optimal establishment of symbiosis (del Giudice *et al.*, 2011). Interestingly, nodules formed by the overexpressing strain (*hmp*$^{++}$) had low NO concentrations and delayed nodule senescence and, conversely, a null mutant strain (*hmp*) produced nodules with high NO concentrations and early senescence (Cam *et al.*, 2012). The Glb of *B. japonicum* has a role in NO detoxification during free-living nitrate-dependent growth (Cabrera *et al.*, 2016), but its role in soybean nodules remains unclear.

There is scarce information on the implication of Glbs in other symbioses. The nodules of *M. truncatula* express two class 3 Glb genes. *Mtru3-1* is predominantly expressed in the infected cells and *Mtru3-2* in the nodule base and vascular tissue. Only *Mtru3-2* is up-regulated in arbuscule-containing cells and in the vascular tissue of mycorrhizal root segments colonized by the fungus *Rhizophagus irregularis* (formerly *Glomus intraradices*) (Vieweg *et al.*, 2005). Very recently, it has been shown that *Slyc1* is up-regulated after infection with *R. irregularis* and that overexpressing or silencing *Slyc1* in the roots deregulates NO concentrations and alters mycorrhization and pathogen infection (Martínez-Medina *et al.*, 2019). These results permit one to conclude that class 1 Glbs, by modulating NO concentrations, are involved in the onset of both the rhizobial and arbuscular mycorrhizal symbioses.

VI. Globins in evolution

Molecular phylogenetic analyses suggest that the 3/3 Glb genes of plants were acquired via horizontal transfer of bacterial flavohemoglobin genes to an ancestral eukaryote, whereas the 2/2 Glb genes were transferred from a bacterium ancestral to Chloroflexi (formerly, 'green non-sulfur bacteria') prior to the evolution of land plants (Vinogradov *et al.*, 2011). For the purposes of this review, we performed detailed phylogenetic analyses of Glbs from algae to

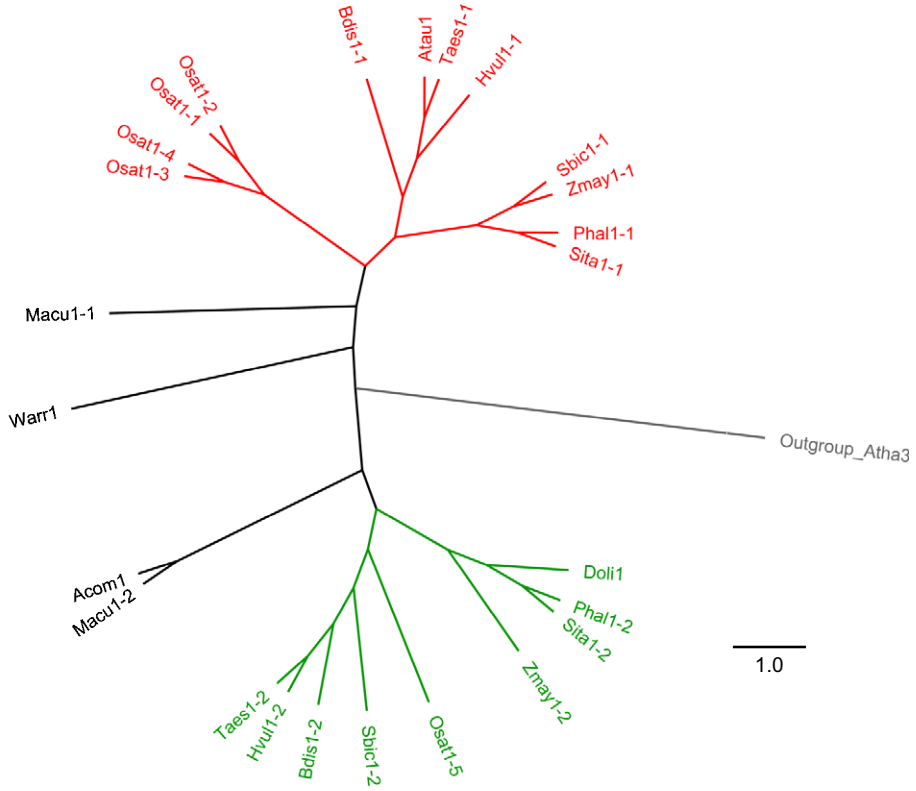

Fig. 4 Phylogenetic tree of class 1 globins of monocots. Note the two clades (colored in red and green) within Poaceae. The sequences outside the two clades belong to other monocot families: Bromeliaceae (Acom1), Araceae (Warr1) and Musaceae (Macu1-1 and Macu1-2). Abbreviations and accession nos are as follows: *Aegilops tauschii* (Atau1, XP_020148046); *Ananas comosus* (Acom1, XP_020091450); *Brachypodium distachyon* (Bdis1-1, Bradi1g69320; Bdis1-2, Bradi2g19690); *Dichanthelium oligosanthes* (Doli1, OEL13199); *Hordeum vulgare* (Hvul1-1, HORVU4Hr1G066200; Hvul1-2, HORVU1Hr1G076460); *Musa acuminata* (Macu1-1, XP_009389441; Macu1-2, GSMUA_Achr5T01820_001); *Oryza sativa* (Osat1-1, Os03g13140; Osat1-2, Os03g12510; Osat1-3, Os03g13150; Osat1-4, Os03g13160; Osat1-5, Os05g44140); *Panicum hallii* (Phal1-1, Pahal.9G547300; Phal1-2, Pahal.3G212300); *Triticum aestivum* (Taes1-1, 4DL_B819BDB3C; Taes1-2, 1AL_762DB7548); *Setaria italica* (Sita1-1, Seita.9G483600; Sita1-2, Seita.3G177600); *Sorghum bicolor* (Sbic1-1, Sobic.001G449600; Sbic1-2, Sobic.009G199601); *Wolffia arrhiza* (Warr1, AEQ39061); *Zea mays* (Zmay1-1, GRMZM2G067402_T02; Zmay1-2, GRMZM2G168898_T01). Outgroup: Arabidopsis (Atha3, At4g32690). Underlined accession nos. correspond to globins whose expression has been examined at the mRNA level and/or those that have been biochemically characterized. The sequence profiles were globally aligned with Clustal Omega (https://www.ebi.ac.uk/Tools/ msa/clustalo/) and trimmed following the protocol of the trimAL software (Capella-Gutiérrez *et al.*, 2009). A maximum likelihood phylogenetic tree using the Subtree Pruning and Regrafting (SPR) method was constructed with PhyML (https://ngphylogeny.fr; Lemoine et al., 2019). The tree and cladogram were midpoint-rooted and plotted with FigTree (http://tree.bio.ed.ac.uk/software/figtree/). The approximate Likelihood-Ratio Test (aLRT) with a seed value of 123 456 and bootstrap analyses with a value of 100 were performed. The bootstrap support values for all branches are 100. aLRT statistics: 0.022 proportion of invariant.

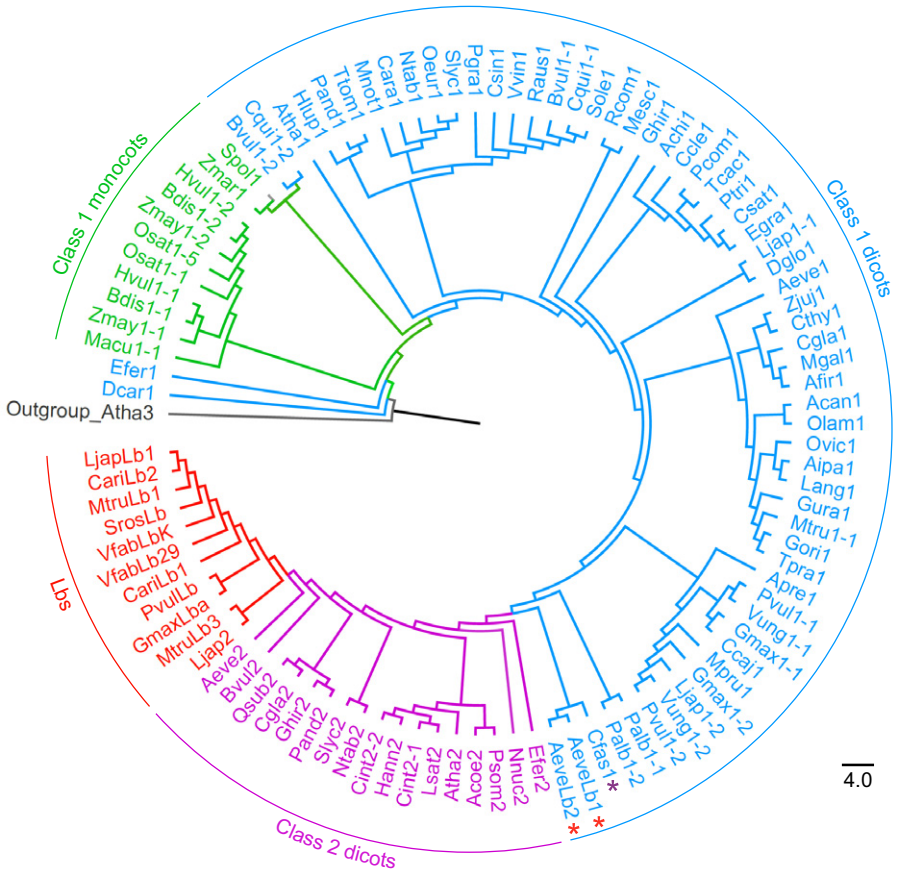

Fig. 5 Phylogenetic tree of class 1 and class 2 globins (Glbs) and leghemoglobins (Lbs) of dicots. Several sequences of monocot Glbs have been included to define the group. Colors are used to denote well-defined groups of class 1 Glbs of monocots (green), class 1 Glbs of dicots (blue), class 2 Glbs of dicots (purple) and Lbs of legumes (red). As expected, Lbs cluster away from class 1 Glbs but close to class 2 Glbs. There is a surprising exception: *Aeschynomene evenia* AeveLb1 and AeveLb2 (*red asterisks*) form a clade, among class 1 Glbs, with *Chamaecrista fasciculata* Cfas1 (*purple asterisk*). These and other data indicate that AeveLbs derive from class 1 Glbs and suggest evolutionary relationships between them and the Glbs of caesalpinioid legumes. For *Casuarina glauca*, we have renamed the original designation of Glb sequences for consistency: Cgla1 clusters with class 1 Glbs, corresponds to previous HbII, and is a nonsymbiotic globin that is expressed in leaves, roots and stems (Christensen *et al.*, 1991); Cgla2 clusters with class 2 Glbs, corresponds to previous HbI, and is a symbiotic globin expressed only in nodules (Kortt *et al.*, 1988). Abbreviations and accession nos. other than those in Fig. 4 are as follows: *Abrus precatorius* (Apre1, XP_027354442); *Actinidia chinensis* (Achi1, PSS19171); *A. evenia* (Aeve1, Ae04g33130; Aeve2, Ae09g18610; AeveLb1, Ae04g33090; AeveLb2, Ae04g33100); *Alnus firma* (Afir1, BAE75956); *Aquilegia coerulea* (Acoe2, Aqcoe2G156900); Arabidopsis (Atha1, At2g16060; Atha2, At3g10520); *Arachis ipaensis* (Aipa1, XP_016187060); *Astragalus canadensis* (Aca1, QAX32747); *Beta vulgaris* (Bvul1-1, AHB20277; Bvul1-2, AHB20278; Bvul2, AHB20279); *Cajanus cajan* (Ccaj1, KYP66039); *Camellia sinensis* (Csin1, XP_028051695); *C. glauca* (Cgla1, CAA37898; Cgla2, AAA33018); *Ceanothus thyrsiflorus* (Cthy1, AZL41245); *C. fasciculata* (Cfas1, ABR68293); *Chenopodium quinoa* (Cqui1-1, XP_021738542; Cqui1-2, XP_021738543); *Cicer arietinum* (CariLb1, Ca_16113; CariLb2, Ca_16084); *Cichorium intybus x endivia* (Cint2-1, CAA07547; Cint2-2, CAB91629); *Citrus clementina* (Ccle1, XP_006423147); *Coffea arabica* (Cara1, XP_027112350); *Cucumis sativus* (Csat1, KGN61772); *Datisca glomerata* (Dglo1, AZL41264); *Daucus carota* (Dcar1, DCAR_028135); *Eucalyptus grandis* (Egra1, XP_010028548); *Euryale ferox* (Efer1, AAQ22728; Efer2, AAQ22729); *Galega orientalis* (Gori1, QAX32755); *Glycine max* (Gmax1-1, NP_001344410; Gmax1-2, ACU23998; GmaxLba, NP_001235928); *Glycyrrhiza uralensis* (Gura1, QAX32717); *Gossypium hirsutum* (Ghir1, AAL09463; Ghir2, AAK21604); *Helianthus annuus* (Hann2, HanXRQChr07g0193521); *Humulus lupulus* (Hlup1, DAA80497); *Lactuca sativa* (Lsat2, XP_023773123); *Lotus japonicus* (Ljap1-1, Lj3g3v3338170; Ljap1-2, Lj3g3v3338180; Ljap2, Lj5g3v1699110; LjapLb1, Lj5g3v0035290.2); *Lupinus angustifolius* (Lang1, OIV99463); *Manihot esculenta* (Mesc1, XP_021599381); *Medicago truncatula* (Mtru1-1, Medtr4g068860; MtruLb1, Medtr5g066070; MtruLb3, Medtr1g090810); *Morus notabilis* (Mnot1, XP_024021257); *Mucuna pruriens* (Mpru1, RDX94852); *Myrica gale* (Mgal1, ABN49927); *Nelumbo nucifera* (Nnuc2, XP_010241647); *Nicotiana tabacum* (Ntab1, XP_016491855; Ntab2, XP_016471234); *Olea europaea* (Oeur1, XP_022864201); *Onobrychis viciifolia* (Ovic1, QAX32726); *Oxytropis lambertii* (Olam1, QAX32736); *Papaver somniferum* (Psom2, XP_026414890); *Parasponia andersonii* (Pand1, AAB86653; Pand2, PON43617); *Phaseolus vulgaris* (Pvul1-1, Phvul.011G048700; Pvul1-2, Phvul.011G048600; PvulLb, Phvul.007G142500); *Populus trichocarpa* (Ptri1, XP_002313074); *Prosopis alba* (Palb1-1, XP_028781526; Palb1-2, XP_028777050); *Punica granatum* (Pgra1, PKI58811); *Pyrus communis* (Pcom1, AAP57677); *Quercus suber* (Qsub2, XP_023876226); *Rheum australe* (Raus1, ACH63214); *Ricinus communis* (Rcom1, XP_002519108); *Sesbania rostrata* (SrosLb, P14848); *Solanum lycopersicum* (Slyc1, NP_001234498; Slyc2, NP_001234111); *Spinacia oleracea* (Sole1, XP_021850850); *Spirodela polyrhiza* (Spol1, Spipo0G0113400); *Theobroma cacao* (Tcac1, XP_007041393); *Trifolium pratense* (Tpra1, PNX93138); *Trema tomentosum* (Ttom1, 3QQQ); *Vicia faba* (VfabLb29, CAA90871; VfabLbK, CAA90869); *Vigna unguiculata* (Vung1-1, XP_027910775; Vung1-2, XP_027910776); *Vitis vinifera* (Vvin1, RVW48423); *Ziziphus jujuba* (Zjuj1, XP_015900815); *Zostera marina* (Zmar1, Zosma33g00440). Outgroup: Arabidopsis (Atha3, At4g32690). Underlined accession nos. correspond to Glbs or Lbs whose expression has been examined and/or those that have been biochemically characterized. The tree was constructed using the same procedures indicated in the legend of Fig. 4, except that it was rooted using Atha3. All branches have bootstrap support values of *c.* 100, except for the branches of Efer1/Dcar1 and the clusters of Palb1-1/Palb1-2 and Ljap1-1/Dglo1, which have values < 70. Approximate Likelihood-Ratio Test statistics: 0.019 proportion of invariant.

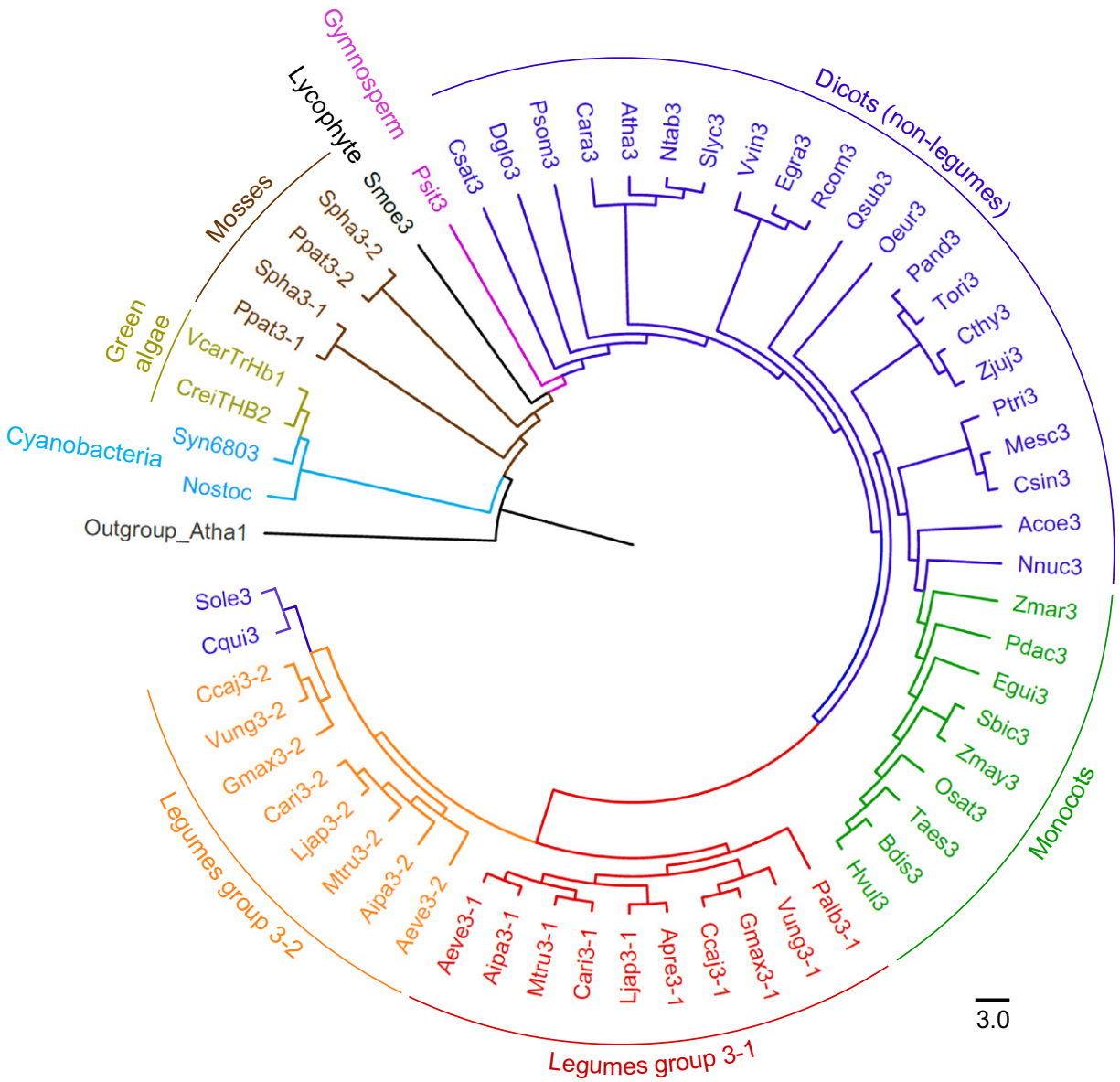

Fig. 6 Phylogenetic tree of truncated or class 3 globins from algae, bryophytes and vascular plants. The sequences of two cyanobacterial TrHbs were included for comparison. For angiosperms, we have colored the lineages of the monocots (green), non-legume dicots (purple) and the two groups of legumes (3-1 in red and 3-2 in orange). Surprisingly, two sequences of Amaranthaceae (spinach Sole3 and quinoa Cqui3) cluster with group 3-2 of legumes. Abbreviations and accession nos. are as follows: *Abrus precatorius* (Apre3-1, XP_027356516); *Aeschynomene evenia* (Aeve3-1, Ae08g12400; Aeve3-2, Ae07g15760); *Aquilegia coerulea* (Acoe3, PIA65070); Arabidopsis (Atha3, At4g32690); *Arachis ipaensis* (Aipa3-1, XP_016182178; Aipa3-2, XP_016162187); *Brachypodium distachyon* (Bdis3, Bradi1g37100); *Cajanus cajan* (Ccaj3-1, XP_020218112; Ccaj3-2, XP_020206657); *Ceanothus thyrsiflorus* (Cthy3, AZL93842); *Chenopodium quinoa* (Cqui3, XP_021746363); *Cicer arietinum* (Cari3-1, XP_004501195; Cari3-2, XP_004498545); *Chlamydomonas reinhardtii* (CreiTHB2, Cre14.g615350); *Citrus sinensis* (Csin3, XP_006476158); *Coffea arabica* (Cara3, XP_027085380); *Cucumis sativus* (Csat3, XP_004138784); *Datisca glomerata* (Dglo3, CAD33536); *Elaeis guineensis* (Egui3, XP_010925331); *Eucalyptus grandis* (Egra3, XP_010049069); *Glycine max* (Gmax3-1, Glyma.04G079200; Gmax3-2, Glyma.14G140400); *Hordeum vulgare* (Hvul3, AAK55410); *Lotus japonicus* (Ljap3-1, Lj1g3v2035270; Ljap3-2, Lj1g3v0948590); *Manihot esculenta* (Mesc3, XP_021606494); *Medicago truncatula* (Mtru3-1, Medtr3g109420; Mtru3-2, Medtr1g008700); *Nelumbo nucifera* (Nnuc3, XP_010249775); *Nicotiana tabacum* (Ntab3, XP_016471167); *Nostoc* sp. (Nostoc, WP_067770579); *Olea europaea* (Oeur3, XP_022854563); *Oryza sativa* (Osat3, Os06g0591600); *Papaver somniferum* (Psom3, XP_026404694); *Parasponia andersonii* (Pand3, PON76481); *Phoenix dactylifera* (Pdac3, XP_008786410); *Physcomitrella patens* (Ppat3-1, XP_024377531; Ppat3-2, XP_024393913); *Picea sitchensis* (Psit3, ABK22150); *Populus trichocarpa* (Ptri3, XP_002309574); *Prosopis alba* (Palb3-1, XP_028802320); *Quercus suber* (Qsub3, XP_023902275); *Ricinus communis* (Rcom3, XP_002516587); *Selaginella moellendorffii* (Smoe3, XP_002991488); *Sphagnum fallax* (Spha3-1, Sphfalx0029s0049; Spha3-2, Sphfalx0052s0112); *Solanum lycopersicum* (Slyc3, XP_004245251); *Sorghum bicolor* (Sbic3, XP_002438623); *Spinacia oleracea* (Sole3, XP_021835671); *Synechocystis* PCC 6803 (Syn6803, pdb 1RTX); *Trema orientale* (Tori3, POO03904); *Triticum aestivum* (Taes3, AAN85433); *Vigna unguiculata* (Vung3-1, XP_027902303; Vung3-2, XP_027938955); *Vitis vinifera* (Vvin3, XP_002284484); *Volvox carteri* (VcarTrHb1, XP_002945683); *Zea mays* (Zmay3, ACG29525); *Ziziphus jujuba* (Zjuj3, XP_015885170); *Zostera marina* (Zmar3, Zosma84g00080). Outgroup: Arabidopsis (Atha1, At2g16060). Underlined accession nos. correspond to globins whose expression has been examined and/or those that have been biochemically characterized. The tree was constructed as indicated in the legend of Fig. 4, except that it was rooted using Atha1. All branches have bootstrap support values of *c.* 100, except for the clades of mosses and Sole3/Cqui3, which have values < 70. Approximate Likelihood-Ratio Test statistics: 0.002 proportion of invariant.

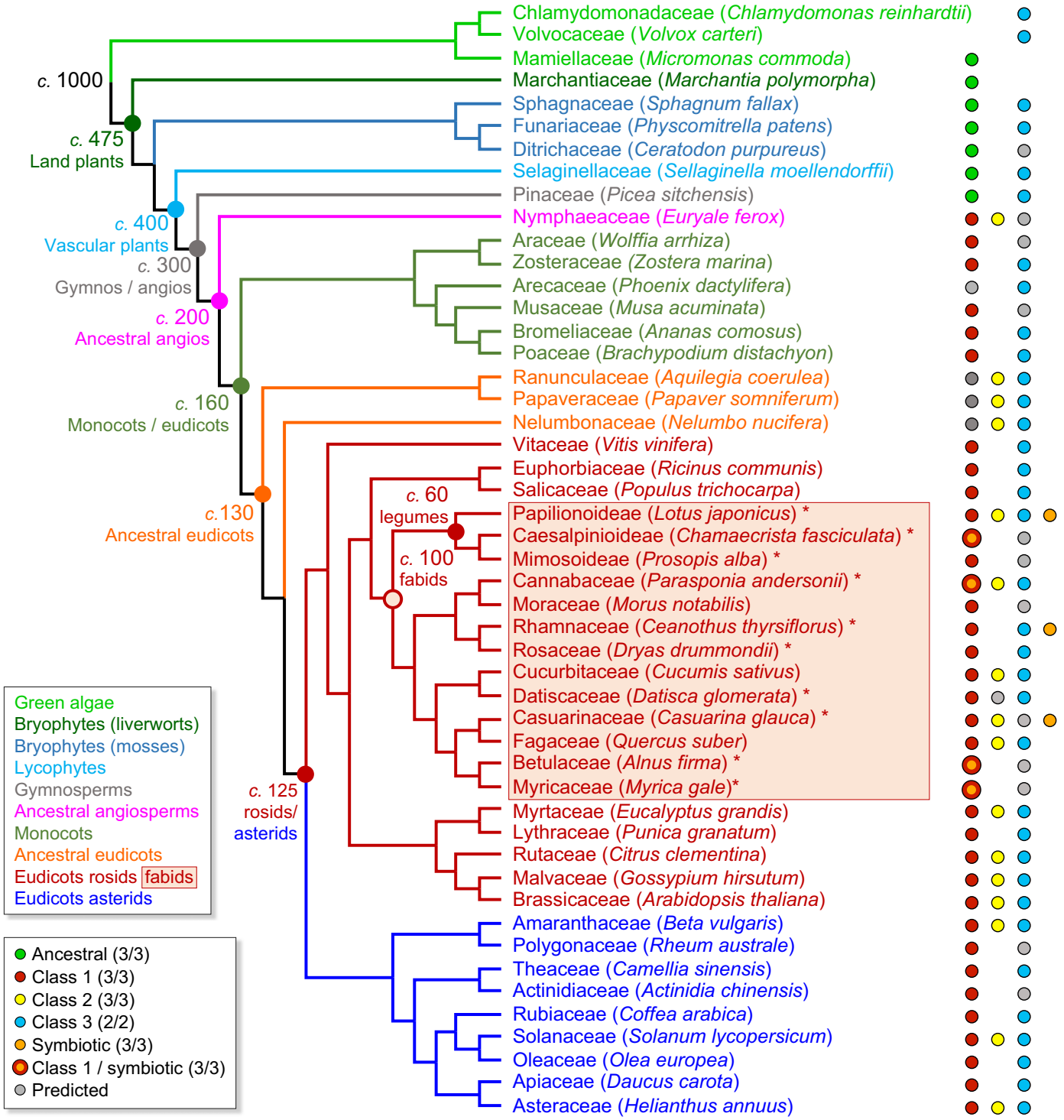

Fig. 7 Evolutionary tree of the plant families examined in this review, with representative species of each family. The phylogeny and the classification system follow those proposed for the different clades of the tree of life of algae and land plants (summarized in Vargas & Zardoya, 2014) and the angiosperms (Angiosperm Phylogeny Group, 2016). Three subfamilies of Leguminosae are shown: caesalpinioids, mimosoids and papilionoids. The tree shows the major divergences during plant evolution in millions of years ago, as estimated by Vargas & Zardoya (2014) for algae and land plants and by Li *et al.* (2019) for gymnosperms and angiosperms. Colored circles indicate that a protein sequence of the corresponding globin (Glb) class is available for at least one species of the plant family. A red/orange concentrical circle indicates that a class 1 Glb also has symbiotic functions. A gray circle indicates that the Glb is likely to be present but its sequence is not available. A colored box indicates the fabids clade, which contains families (marked with an asterisk) forming rhizobial or actinorhizal nodules (the so-called 'nitrogen-fixing clade'). The Rosaceae includes only a few nodulating genera and species. For example, *Dryas drummondii* is a nodulating species, whereas *Dryas octopetala* is not (Billault-Penneteau *et al.*, 2019). The Coriariaceae and Elaeagnaceae are also included in this clade, but no Glb sequences are yet available.

eudicots, encompassing a total of 47 families, 88 species and 181 Glb sequences. The analysis of class 1 or 2 Glbs was performed with Atha3 as outgroup and the analysis of class 3 Glbs with Atha1 as outgroup, which highlights the clear evolutionary separation between class 1 or 2 Glbs and class 3 Glbs. Fig. 4 shows a phylogenetic tree of monocots with two distinct clades. Clade 1 contains some Glbs (Hvul1-1, Osat1-1 and Zmay1-1) for which biochemical studies are available, whereas clade 2 contains so far unexplored Glbs. Both clades of monocot Glbs belong to class 1, as can be inferred from a phylogenetic analysis of class 1 and 2 Glbs of angiosperms shown in Fig. 5. This cladogram also evidences a clear separation of the two classes of Glbs in dicotyledonous plants and provides interesting information about legume Glbs. Lbs cluster away from class 1 Glbs but close to class 2 Glbs, which was expected because Lbs derive from class 2 Glbs. However, there is an exception because Lbs from the tropical legume *Aeschynomene evenia* (AeveLb1 and AeveLb2, *red asterisks*) derive from class 1 Glbs, as can be observed in the cladogram and was confirmed by biochemical characterization (J-F. Arrighi *et al.*, unpublished). Notably, these two Lbs form a clade with the Glb of the primitive caesalpinioid legume *Chamaecrista fasciculata* (Cfas1, *purple asterisk*), suggesting an evolutionary relationship.

Fig. 6 is a phylogenetic tree of class 3 Glbs from cyanobacteria, green algae, mosses, lycophytes, gymnosperms and angiosperms, showing an increasing distance from the outgroup. Within bryophytes, it is puzzling that *P. patens* and *S. phallax* have two class 3 Glbs, whereas *M. polymorpha* has none (Section III; Figs 6 and 7). Within angiosperms, we found no obvious patterns because class 3 Glbs of monocots are placed in between those of non-leguminous and leguminous plants. Notably, two groups of class 3 Glbs can be observed in legumes, but there are no molecular or functional clues that could explain them other than differences in the expression profiles, which suggests that clade 3-1 Glbs have a role in nodulation (Section IV.4).

Fig. 7 shows a phylogenetic tree of the plant families that have been examined in this review with the Glbs found so far within each family, although not necessarily present in the representative species. In the tree we have marked the major transitions that occurred during plant evolution. Combining the evolutionary divergences with Glb compositions, some observations are worth mentioning. First, the appearance of land plants at *c.* 475 Ma can be associated with the appearance of an ancestral 3/3 Glb in bryophytes and pteridophytes sharing 62–69% similarity with class 1 Glbs of angiosperms (Section III; Fig. 3). Interestingly, the green algae *M. commoda* and *O. tauri* also have 3/3 Glbs with 52–55% similarity with class 1 Glbs, but the corresponding genes have no introns. Second, the split of spermatophytes (seed plants) into gymnosperms and angiosperms occurred at *c.* 300 Ma, and the only known 3/3 Glb of a gymnosperm, Psit, shares 75–78% similarity with class 1 Glbs of angiosperms. This relatively high sequence homology strongly suggests Psit is a class 1 Glb, but its biochemical characterization is required (Section III; Fig. 3). Third, ancestral angiosperms (Nymphaeaceae) already had class 1 and 2 Glbs at *c.* 200 Ma. Because the diversification of angiosperms was preceded by a whole-genome duplication event, it is likely that this duplication marks the point at which the two Glb classes emerged from the ancestral 3/3 Glb (Vázquez-Limón *et al.*, 2012b). Fourth, the divergence between monocots and eudicots occurred at *c.* 160 Ma. Because monocots have only class 1 Glbs, they probably lost class 2 Glbs (Rodríguez-Alonso & Arredondo-Peter, 2012). Finally, the appearance of the fabids clade *c.* 100 Ma would mark the origin of symbiotic Glbs and later on the appearance of Leguminosae and hence of Lbs at *c.* 60 Ma. The convergent evolution of O_2 transport in plant Glbs and Lbs has been described in detail elsewhere (Kakar *et al.*, 2010; Sturms *et al.*, 2010).

Table 1 Important research targets for elucidating structure and function of plant globins.

Globin	Objectives for future research
Glb1	For monocots, structural and functional characterization of clade 2 Glbs and comparison with their clade 1 counterparts
Glb2	For dicots, structural and functional studies, including crystallization, of relevant class 2 Glbs, such as Pand2
Glb3	For legumes, structural and functional differences between the two groups of class 3 Glbs
Glb	Structure and function of globins from algae, bryophytes, lycophytes, gymnosperms and ancestral angiosperms
Glb	Function of Glbs in cytoplasm and chloroplasts; interactions with other proteins
Glb, Lb	Mechanism(s), in particular potential reductases, for reduction of ferric Glbs and Lbs
Glb, Lb	Function of Glbs and Lbs in nuclei; interactions with nuclear proteins; interaction of Glb and Lb gene promoters with transcription factors
Glb, Lb	Correlation (or lack of) between protein abundances and transcript levels. Post-translational modifications?
Lb	Specific function of Lb isoproteins in model legumes (twelve in *Medicago*, three in *Lotus*)
Lb	Role of reactive oxygen species, heme oxygenase and proteases in Lb degradation
Lb	Function of Lbs in homeostasis of reactive oxygen and nitrogen species

VII. Perspectives for future research

Despite 20 years having passed since their discovery, many aspects of the structures and especially the functions of Glbs remain enigmatic. While significant details are known about their biophysical and biochemical properties *in vitro*, progress on their physiological roles has been harder to come by. In this review, we have identified important gaps in our knowledge of the structures and functions of plant Glbs. Some of the research targets are listed in Table 1. The structures and functions of Glbs belonging to primitive plants and to some clades of monocots and legumes are unknown. Transcript data, when reported or culled from databases, indicate peak expression of specific Glbs at specific developmental stages or stressful conditions and in specific tissues, indicating tight spatiotemporal regulation. Although strong evidence has been accrued for a role of class 1 Glbs in NO homeostasis and tolerance to hypoxia, the involvement of those proteins in a NOD-like reaction requires a cognate reductase that has yet to be identified. Also, Glb functions unrelated to NO are expected to be discovered because remarkable differences exist in the structures of the various classes and even between the members of each class. Last but not least, the function of the individual Lbs

present in the cytoplasm and nuclei of nodule cells, their regulation at the mRNA, protein and post-translational levels, as well as their possible interactions *in vivo* with reactive oxygen and nitrogen species, deserve thorough investigation.

Acknowledgements

We thank Carmen Pérez-Rontomé for help with some of the figures and Jean-François Arrighi, Katharina Pawlowski, Maria Jesús Delgado, René Geurts and Raúl Arredondo-Peter for sharing information. We are also grateful to two anonymous reviewers for helpful comments. Part of the work described in the lab of MB was funded by the Spanish Ministry of Science and Innovation-European Regional Development Fund (grant AGL2017-85775-R) and by Government of Aragón, Spain (group A09_17R).

ORCID

Manuel Becana https://orcid.org/0000-0002-1083-0804
Pilar Catalán https://orcid.org/0000-0001-7793-5259
Mark S. Hargrove https://orcid.org/0000-0002-7127-3183
Gautam Sarath https://orcid.org/0000-0002-3145-9493
Inmaculada Yruela https://orcid.org/0000-0003-3608-4720

References

Alagurajan J, Singh Athwal N, Hargrove MS. 2018. Steady-state kinetics of phytoglobin-catalyzed reduction of hydroxylamine to ammonium. *Biochemistry* **57**: 4824–4832.

Alayash AI. 2001. Oxidative mechanisms of hemoglobin-based blood substitutes. *Artificial Cells, Blood Substitutes, and Biotechnology* **29**: 415–425.

Andersson CR, Jensen EO, Llewellyn DJ, Dennis ES, Peacock WJ. 1996. A new hemoglobin gene from soybean: a role for hemoglobin in all plants. *Proceedings of the National Academy of Sciences, USA* **93**: 5682–5687.

Angiosperm Phylogeny Group. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society* **181**: 1–20.

Appleby CA. 1984. Leghemoglobin and *Rhizobium* respiration. *Annual Review of Plant Physiology* **35**: 443–478.

Appleby CA, Bogusz D, Dennis ES, Peacock WJ. 1988. A role for haemoglobin in all plant roots? *Plant, Cell & Environment* **11**: 359–367.

Appleby CA, Tjepkema JD, Trinick MJ. 1983. Hemoglobin in a nonleguminous plant, *Parasponia*: possible origin and function in nitrogen fixation. *Science* **220**: 951–953.

Arredondo-Peter R, Hargrove MS, Moran JF, Sarath G, Klucas RV. 1998. Plant hemoglobins. *Plant Physiology* **118**: 1121–1125.

Arredondo-Peter R, Hargrove MS, Sarath G, Moran JF, Lohrman J, Olson JS, Klucas RV. 1997. Rice hemoglobins. Gene cloning, analysis, and O_2-binding kinetics of a recombinant protein synthesized in *Escherichia coli*. *Plant Physiology* **115**: 1259–1266.

Athwal NS, Alagurajan J, Andreotti AH, Hargrove MS. 2016. Role of reversible histidine coordination in hydroxylamine reduction by plant hemoglobins (phytoglobins). *Biochemistry* **55**: 5809–5817.

Athwal NS, Alagurajan J, Sturms R, Fulton DB, Andreotti AH, Hargrove MS. 2015. Electron self-exchange in hemoglobins revealed by deutero-hemin substitution. *Journal of Inorganic Biochemistry* **150**: 139–147.

Becana M, Klucas RV. 1990. Enzymatic and nonenzymatic mechanisms for ferric leghemoglobin reduction in legume nodules. *Proceedings of the National Academy of Sciences, USA* **87**: 7295–7299.

Berger A, Guinand S, Boscari A, Puppo A, Brouquisse R. 2020. *Medicago truncatula* phytoglobin 1.1 controls symbiotic nodulation and nitrogen fixation via the regulation of nitric oxide level. *New Phytologist* **227**: 84–98.

Billault-Penneteau B, Sandré A, Folgmann J, Parniske M, Pawlowski K. 2019. Dryas as a model for studying the root symbioses of the Rosaceae. *Frontiers in Plant Science* **10**: 661.

Bruno S, Faggiano S, Spyrakis F, Mozzarelli A, Cacciatori E, Dominici P, Grandi E, Abbruzzetti S, Viappiani C. 2007. Different roles of protein dynamics and ligand migration in non-symbiotic hemoglobins AHb1 and AHb2 from *Arabidopsis thaliana*. *Gene* **398**: 224–233.

Burmester T, Hankeln T. 2014. Function and evolution of vertebrate globins. *Acta Physiologica* **211**: 501–514.

Bustamante JP, Radusky L, Boechi L, Estrin DA, ten Have A, Martí MA. 2016. Evolutionary and functional relationships in the truncated hemoglobin family. *PLoS Computational Biology* **12**: e1004701.

Bustos-Díaz ED, Barona-Gómez F, Cibrián-Jaramillo A. 2019. Cyanobacteria in nitrogen-fixing symbioses. In: Mishra AK, Tiwari DN, Rai AN, eds. *Cyanobacteria. From basic science to applications*. London, UK: Academic Press, 29–42.

Bustos-Sanmamed P, Tovar-Méndez A, Crespi M, Sato S, Tabata S, Becana M. 2011. Regulation of non-symbiotic and truncated hemoglobin genes of *Lotus japonicus* in plant organs and in response to nitric oxide and hormones. *New Phytologist* **189**: 765–776.

Cabrera JJ, Salas A, Torres MJ, Bedmar EJ, Richardson DJ, Gates AJ, Delgado MJ. 2016. An integrated biochemical system for nitrate assimilation and nitric oxide detoxification in *Bradyrhizobium japonicum*. *Biochemical Journal* **473**: 297–309.

Calvo-Begueria L, Cuypers B, Van Doorslaer S, Abbruzzetti S, Bruno S, Berghmans H, Dewilde S, Ramos J, Viappiani C, Becana M. 2017. Characterization of the heme pocket structure and ligand binding kinetics of non-symbiotic hemoglobins from the model legume *Lotus japonicus*. *Frontiers in Plant Science* **8**: 407.

Calvo-Begueria L, Rubio MC, Martínez JI, Pérez-Rontomé C, Delgado MJ, Bedmar EJ, Becana M. 2018. Redefining nitric oxide production in legume nodules through complementary insights from electron paramagnetic resonance spectroscopy and specific fluorescent probes. *Journal of Experimental Botany* **69**: 3703–3714.

Cam Y, Pierre O, Boncompagni E, Hérouart D, Meilhoc E, Bruand C. 2012. Nitric oxide (NO): a key player in the senescence of *Medicago truncatula* root nodules. *New Phytologist* **196**: 548–560.

Capella-Gutiérrez S, Silla-Martinez JM, Gabaldòn T. 2009. TrimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**: 1972–1973.

Chamizo-Ampudia A, Sanz-Luque E, Llamas A, Galvan A, Fernandez E. 2017. Nitrate reductase regulates plant nitric oxide homeostasis. *Trends in Plant Science* **22**: 163–174.

Christensen T, Dennis ES, Peacock JW, Landsmann J, Marcker JA. 1991. Hemoglobin genes in non-legumes: cloning and characterization of a *Casuarina glauca* hemoglobin gene. *Plant Molecular Biology* **16**: 339–344.

Coats V, Schwintzer CR, Tjepkema JD. 2009. Truncated hemoglobins in *Frankia* CcI3: effects of nitrogen source, oxygen concentration, and nitric oxide. *Canadian Journal of Microbiology* **55**: 867–873.

Couture M, Chamberland H, St-Pierre B, Lafontaine J, Guertin M. 1994. Nuclear genes encoding chloroplast hemoglobins in the unicellular green alga *Chlamydomonas eugametos*. *Molecular General Genetics* **243**: 185–197.

Couture M, Guertin M. 1996. Purification and spectroscopic characterization of a recombinant chloroplastic hemoglobin from the green unicellular alga *Chlamydomonas eugametos*. *European Journal of Biochemistry* **242**: 779–787.

Das TK, Lee HC, Duff SMG, Hill RD, Peisach J, Rousseau DL, Wittenberg BA, Wittenberg JB. 1999. The heme environment of barley hemoglobin. *Journal of Biological Chemistry* **274**: 4207–4212.

De Henau S, Tilleman L, Vangheel M, Luyckx E, Trashin S, Pauwels M, Germani F, Vlaeminck C, Vanfleteren JR, Bert W *et al.* 2015. A redox signalling globin is essential for reproduction in *Caenorhabditis elegans*. *Nature Communications* **6**: 8782.

Duff SM, Wittenberg JB, Hill RD. 1997. Expression, purification, and properties of recombinant barley (*Hordeum* sp.) hemoglobin. Optical spectra and reactions with gaseous ligands. *Journal of Biological Chemistry* **272**: 16746–16752.

Elhiti M, Hebelstrup KH, Wang A, Li C, Cui Y, Hill RD, Stasolla C. 2013. Function of type-2 Arabidopsis hemoglobin in the auxin-mediated formation of embryogenic cells during morphogenesis. *The Plant Journal* **74**: 946–958.

Fernández I, Vinogradov SN, Arredondo-Peter R. 2010. Identification and *in silico* characterization of a putative ancestor to land plant non-symbiotic hemoglobins from the prasinophyceae algae *Micromonas* and *Ostreococcus*. *Global Journal of Biochemistry* **1**: 18–30.

Fleming AI, Wittenberg JB, Wittenberg BA, Dudman WF, Appleby CA. 1987. The purification, characterization and ligand-binding kinetics of hemoglobins from root nodules of the non-leguminous *Casuarina glauca* - *Frankia* symbiosis. *Biochimica et Biophysica Acta* **91**: 209–220.

Fukudome M, Calvo-Beguería L, Kado T, Osuki K, Rubio MC, Murakami E, Nagata M, Kucho K, Sandal N, Stougaard J *et al.* 2016. Hemoglobin LjGlb1-1 is involved in nodulation and regulates the level of nitric oxide in the *Lotus japonicus*-*Mesorhizobium loti* symbiosis. *Journal of Experimental Botany* **67**: 5275–5283.

Fukudome M, Watanabe E, Osuki K, Imaizumi R, Aoki T, Kucho K, Becana M, Uchiumi T. 2019. Stably transformed *Lotus japonicus* plants overexpressing phytoglobin LjGlb1-1 show decreased nitric oxide levels in roots and nodules as well as delayed nodule senescence. *Plant and Cell Physiology* **60**: 816–825.

Gardner PR. 2012. Hemoglobin: a nitric-oxide dioxygenase. *Scientifica* **2012**: 683729.

Gardner PR, Gardner AM, Martin LA, Salzman AL. 1998. Nitric oxide dioxygenase: an enzymic function for flavohemoglobin. *Proceedings of the National Academy of Sciences, USA* **95**: 10378–10383.

Garrocho-Villegas V, Arredondo-Peter R. 2008. Molecular cloning and characterization of a moss (*Ceratodon purpureus*) nonsymbiotic hemoglobin provides insight into the early evolution of plant nonsymbiotic hemoglobins. *Molecular Biology and Evolution* **25**: 1482–1487.

Garrocho-Villegas V, Gopalasubramaniam SK, Arredondo-Peter R. 2007. Plant hemoglobins: what we know six decades after their discovery? *Gene* **398**: 78–85.

del Giudice J, Cam Y, Damiani I, Fung-Chat F, Meilhoc E, Bruand C, Brouquisse R, Puppo A, Boscari A. 2011. Nitric oxide is required for an optimal establishment of the *Medicago truncatula–Sinorhizobium meliloti* symbiosis. *New Phytologist* **191**: 405–417.

Grossman AR, Karpowicz SJ, Heinnickel M, Dewez D, Hamel B, Dent R, Niyogi KK, Johnson X, Alric J, Wollman FA *et al.* 2010. Phylogenomic analysis of the Chlamydomonas genome unmasks proteins potentially involved in photosynthetic function and regulation. *Photosynthesis Research* **106**: 3–17.

Gupta KJ, Igamberdiev AU. 2016. Reactive nitrogen species in mitochondria and their implications in plant energy status and hypoxic stress tolerance. *Frontiers in Plant Science* **7**: 369.

Halder P, Trent JT III, Hargrove MS. 2007. Influence of the protein matrix on intramolecular histidine ligation in ferric and ferrous hexacoordinate hemoglobins. *Proteins* **66**: 172–182.

Hargrove MS, Brucker EA, Stec B, Sarath G, Arredondo-Peter R, Klucas RV, Olson JS, Phillips GN Jr. 2000. Crystal structure of a nonsymbiotic plant hemoglobin. *Structure* **8**: 1005–1014.

Hartman S, Liu Z, van Veen H, Vicente J, Reinen E, Martopawiro S, Zhang H, van Dongen N, Bosman F, Bassel GW *et al.* 2019. Ethylene-mediated nitric oxide depletion pre-adapts plants to hypoxia stress. *Nature Communications* **10**: 4020.

Hebelstrup KH, Jensen EO. 2008. Expression of NO scavenging hemoglobin is involved in the timing of bolting in *Arabidopsis thaliana*. *Planta* **227**: 917–927.

Hebelstrup KH, Jensen EO, Hill RD. 2008. Bioimaging techniques for subcellular localization of plant hemoglobins and measurement of hemoglobin-dependent nitric oxide scavenging in planta. *Methods in Enzymology* **437**: 595–604.

Hebelstrup KH, Shah JK, Igamberdiev AU. 2013. The role of nitric oxide and hemoglobin in plant development and morphogenesis. *Physiologia Plantarum* **148**: 457–469.

Hemschemeier A, Düner M, Casero D, Merchant SS, Winkler M, Happe T. 2013. Hypoxic survival requires a 2-on-2 hemoglobin in a process involving nitric oxide. *Proceedings of the National Academy of Sciences, USA* **110**: 10854–10859.

Herold S, Puppo A. 2005. Oxyleghemoglobin scavenges nitrogen monoxide and peroxynitrite: a possible role in functioning nodules? *Journal of Biological Inorganic Chemistry* **10**: 935–945.

Hill RD. 2012. Non-symbiotic haemoglobins - What's happening beyond nitric oxide scavenging? *AoB Plants* **2012**: pls004.

Hill RD, Hargrove MS, Arredondo-Peter R. 2016. Phytoglobin: a novel nomenclature for plant globins accepted by the globin community at the 2014 XVIII conference on Oxygen-Binding and Sensing Proteins. *F1000Research* **5**: 212.

Hoy JA, Hargrove MS. 2008. The structure and function of plant hemoglobins. *Plant Physiology and Biochemistry* **46**: 371–379.

Hoy JA, Robinson H, Trent JT III, Kakar S, Smagghe BJ, Hargrove MS. 2007a. Plant hemoglobins: a molecular fossil record for the evolution of oxygen transport. *Journal of Molecular Biology* **371**: 168–179.

Hoy JA, Smagghe BJ, Halder P, Hargrove MS. 2007b. Covalent heme attachment in *Synechocystis* hemoglobin is required to prevent ferrous heme dissociation. *Protein Science* **16**: 250–260.

Huang S, Hill RD, Wally OS, Dionisio G, Ayele BT, Jami SK, Stasolla C. 2014. Hemoglobin control of cell survival/death decision regulates *in vitro* plant embryogenesis. *Plant Physiology* **165**: 810–825.

Hunt PW, Klok EJ, Trevaskis B, Watts RA, Ellis MH, Peacock WJ, Dennis ES. 2002. Increased level of hemoglobin 1 enhances survival of hypoxic stress and promotes early growth in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences, USA* **99**: 17197–17202.

Hunt PW, Watts RA, Trevaskis B, Llewelyn DJ, Burnell J, Dennis ES, Peacock WJ. 2001. Expression and evolution of functionally distinct haemoglobin genes in plants. *Plant Molecular Biology* **47**: 677–692.

Igamberdiev AU, Bykova NV, Hill RD. 2006. Nitric oxide scavenging by barley hemoglobin is facilitated by a monodehydroascorbate reductase-mediated ascorbate reduction of methemoglobin. *Planta* **223**: 1033–1040.

Igamberdiev AU, Hill RD. 2004. Nitrate, NO and haemoglobin in plant adaptation to hypoxia: an alternative to classic fermentation pathways. *Journal of Experimental Botany* **55**: 2473–2482.

Jasid S, Simontacchi S, Bartoli CG, Puntarulo S. 2006. Chloroplasts as a nitric oxide cellular source. Effect of reactive nitrogen species on chloroplastic lipids and proteins. *Plant Physiology* **142**: 1246–1255.

Ji L, Wood S, Becana M, Klucas RV. 1991. Purification and characterization of soybean ferric leghemoglobin reductase. *Plant Physiology* **96**: 32–37.

Johnson EA, Lecomte JTJ. 2013. The globins of cyanobacteria and algae. *Advances in Microbial Physiology* **63**: 195–272.

Johnson EA, Rice SL, Preimesberger MR, Nye DB, Gilevicius L, Wenke BB, Brown JM, Witman GB, Lecomte JTJ. 2014. Characterization of THB1, a *Chlamydomonas reinhardtii* truncated hemoglobin: linkage to nitrogen metabolism and identification of lysine as the distal heme ligand. *Biochemistry* **53**: 4573–4589.

Jokipii-Lukkari S, Kastaniotis AJ, Parkash V, Sundström R, Leiva-Eriksson N, Nymalm Y, Blokhina O, Kukkola E, Fagerstedt KV, Salminen TA *et al.* 2016. Dual targeted poplar ferredoxin NADP$^+$ oxidoreductase interacts with hemoglobin 1. *Plant Science* **247**: 138–149.

Jun H-K, Sarath G, Moran JF, Becana M, Klucas RV, Wagner FW. 1994a. Characteristics of modified leghemoglobins isolated from soybean (*Glycine max* Merr.) root nodules. *Plant Physiology* **104**: 1231–1236.

Jun H-K, Sarath G, Wagner FW. 1994b. Detection and purification of modified leghemoglobins from soybean root nodules. *Plant Science* **100**: 31–40.

Kakar S, Hoffman FG, Storz JF, Fabian M, Hargrove MS. 2010. Structure and reactivity of hexacoordinate hemoglobins. *Biophysical Chemistry* **152**: 1–14.

Kakar S, Sturms R, Tiffany A, Nix JC, DiSpirito AA, Hargrove MS. 2011. Crystal structures of *Parasponia* and *Trema* hemoglobins: differential heme coordination is linked to quaternary structure. *Biochemistry* **50**: 4273–4280.

Kapoor K, Mira MM, Ayele BT, Nguyen TN, Hill RD, Stasolla C. 2018. Phytoglobins regulate nitric-oxide dependent abscisic acid synthesis and ethylene-induced programmed cell death in developing maize somatic embryos. *Planta* **247**: 1277–1291.

Kim DY, Hong MJ, Lee YJ, Lee MB, Seo YW. 2013. Wheat truncated hemoglobin interacts with photosystem I PSK-I subunit and photosystem II subunit PsbS1. *Biologia Plantarum* **57**: 281–290.

Kortt AA, Trinick MJ, Appleby CA. 1988. Amino acid sequences of hemoglobins I and II from root nodules of the non-leguminous *Parasponia rigida-rhizobium* symbiosis, and a correction of the sequence of hemoglobin I from *Parasponia andersonii*. *European Journal of Biochemistry* **175**: 141–149.

Kuruthukulangarakoola GT, Zhang J, Albert A, Winkler B, Lang H, Buegger F, Gaupels F, Heller W, Michalke B, Sarioglu H *et al.* 2017. Nitric oxide-fixation

by non-symbiotic haemoglobin proteins in *Arabidopsis thaliana* under N-limited conditions. *Plant, Cell & Environment* **40**: 36–50.

Landsman J, Dennis ES, Higgins TJV, Appeby CA, Kortt AA, Peacock WJ. 1986. Common evolutionary origin of legume and non-legume plant haemoglobins. *Nature* **324**: 166–168.

Lee KK, Shearman LL, Erickson BK, Klucas RV. 1995. Ferric leghemoglobin in plant-attached leguminous nodules. *Plant Physiology* **109**: 261–267.

Leiva-Eriksson N, Pin PA, Kraft T, Dohm JC, Minoche AE, Himmelbauer H, Bülow L. 2014. Differential expression patterns of non-symbiotic hemoglobins in sugar beet (*Beta vulgaris* ssp. *vulgaris*). *Plant and Cell Physiology* **55**: 834–844.

Leiva-Eriksson N, Reeder BJ, Wilson MT, Bülow L. 2019. Sugar beet hemoglobins: reactions with nitric oxide and nitrite reveal differential roles for nitrogen metabolism. *Biochemical Journal* **476**: 2111–2125.

Lemoine F, Correia D, Lefort V, Doppelt-Azeroual O, Mareuil F, Cohen-Boulakia S, Gascuel O. 2019. NGPhylogeny.fr: new generation phylogenetic services for non-specialists. *Nucleic Acids Research* **47**: W260–W265.

Li HT, Yi TS, Gao LM, Ma PF, Zhang T, Yang JB, Gitzendanner MA, Fritsch PW, Cai J, Luo Y *et al.* 2019. Origin of angiosperms and the puzzle of the Jurassic gap. *Nature Plants* **5**: 461–470.

Liu L, Zeng M, Hausladen A, Heitman J, Stamler JS. 2000. Protection from nitrosative stress by yeast flavohemoglobin. *Proceedings of the National Academy of Sciences, USA* **97**: 4672–4676.

Martínez-Medina A, Pescador L, Fernández I, Rodríguez-Serrano M, García JM, Romero-Puertas MC, Pozo MJ. 2019. Nitric oxide and phytoglobin PHYTOGB1 are regulatory elements in the *Solanum lycopersicum–Rhizophagus irregularis* mycorrhizal symbiosis. *New Phytologist* **223**: 1560–1574.

Mathieu C, Moreau S, Frendo P, Puppo A, Davies MJ. 1998. Direct detection of radicals in intact soybean nodules: presence of nitric oxide-leghemoglobin complexes. *Free Radical Biology and Medicine* **24**: 1242–1249.

Matilla AJ, Rodríguez-Gacio MC. 2013. Non-symbiotic hemoglobins in the life of seeds. *Phytochemistry* **87**: 7–15.

Mira MM, Wally OSD, Elhiti M, El-Shanshory A, Reddy DS, Hill RD, Stasolla C. 2016. Jasmonic acid is a downstream component in the modulation of somatic embryogenesis by Arabidopsis class 2 phytoglobin. *Journal of Experimental Botany* **67**: 2231–2246.

Mot AC, Puscas C, Miclea P, Naumova-Letia G, Dorneanu S, Podar D, Dissmeyer N, Silaghi-Dumitrescu R. 2018. Redox control and autoxidation of class 1, 2 and 3 phytoglobins from *Arabidopsis thaliana*. *Scientific Reports* **8**: 13714.

Mukhi N, Dhindwal S, Uppal S, Kapoor A, Arya R, Kumar P, Kaur J, Kundu S. 2016. Structural and functional significance of the N- and C-terminal appendages in *Arabidopsis* truncated hemoglobin. *Biochemistry* **55**: 1724–1740.

Mukhi N, Dhindwal S, Uppal S, Kumar P, Kaur J, Kundu S. 2013. X-ray crystallographic structural characteristics of *Arabidopsis* hemoglobin I and their functional implications. *Biochimica et Biophysica Acta* **1834**: 1944–1956.

Mur LAJ, Sivakumaran A, Mandon J, Cristescu SM, Harren FJM, Hebelstrup KH. 2012. Haemoglobin modulates salicylate and jasmonate/ethylene-mediated resistance mechanisms against pathogens. *Journal of Experimental Botany* **63**: 4375–4387.

Nagata M, Murakami E, Shimoda Y, Shimoda-Sasakura F, Kucho K, Suzuki A, Abe M, Higashi S, Uchiumi T. 2008. Expression of a class 1 hemoglobin gene and production of nitric oxide in response to symbiotic and pathogenic bacteria in *Lotus japonicus*. *Molecular Plant–Microbe Interactions* **21**: 1175–1183.

Navascués J, Pérez-Rontomé C, Gay M, Marcos M, Yang F, Walker FA, Desbois A, Abián J, Becana M. 2012. Leghemoglobin green derivatives with nitrated hemes evidence production of highly reactive nitrogen species during aging of legume nodules. *Proceedings of the National Academy of Sciences, USA* **109**: 2660–2665.

Niemann J, Tisa LS. 2008. Nitric oxide and oxygen regulate truncated hemoglobin gene expression in *Frankia* strain CcI3. *Journal of Bacteriology* **190**: 7864–7867.

Oliveira HC, Salgado I, Sodek L. 2013. Involvement of nitrite in the nitrate-mediated modulation of fermentative metabolism and nitric oxide production of soybean roots during hypoxia. *Planta* **237**: 255–264.

Ott T, van Dongen JT, Günther C, Krusell L, Desbrosses G, Vigeolas H, Bock V, Czechowski T, Geigenberg P, Udvardi M. 2005. Symbiotic leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth and development. *Current Biology* **15**: 531–535.

Ott T, Sullivan J, James EK, Flemetakis E, Günther C, Gibon Y, Ronson C, Udvardi M. 2009. Absence of symbiotic leghemoglobins alters bacteroid and plant cell differentiation during development of *Lotus japonicus* root nodules. *Molecular Plant–Microbe Interactions* **22**: 800–808.

Pathirana SM, Tjepkema JD. 1995. Purification of hemoglobin from the actinorhizal root nodules of *Myrica gale* L. *Plant Physiology* **107**: 827–831.

Pawlowski K, Jacobsen KR, Alloisio N, Denison RF, Klein M, Tjepkema JD, Winzer T, Sirrenberg A, Guan C, Berry AM. 2007. Truncated hemoglobins in actinorhizal nodules of *Datisca glomerata*. *Plant Biology* **9**: 776–785.

Perazzolli M, Dominici P, Romero-Puertas MC, Zago E, Zeier J, Sonoda M, Lamb C, Delledonne M. 2004. Arabidopsis nonsymbiotic hemoglobin AHb1 modulates nitric oxide bioactivity. *Plant Cell* **16**: 2785–2794.

Potts M, Angeloni SV, Ebel RE, Bassam D. 1992. Myoglobin in a cyanobacterium. *Science* **256**: 1690–1691.

Preimesberger MR, Majumdar A, Lecomte JTJ. 2017. Dynamics of lysine as a heme axial ligand: NMR analysis of the *Chlamydomonas reinhardtii* hemoglobin THB1. *Biochemistry* **56**: 551–569.

Reeder BJ, Hough MA. 2014. The structure of a class 3 nonsymbiotic plant haemoglobin from *Arabidopsis thaliana* reveals a novel N-terminal helical extension. *Acta Crystallographica D: Biological Crystallography* **70**: 1411–1418.

Rodríguez-Alonso G, Arredondo-Peter R. 2012. Phylogenetic analysis reveals an apparent duplication of the non-symbiotic hemoglobin 1 gene early in the evolution of monocotyledonous plants. *ScienceJet* **1**: 27.

Ross EJH, Shearman L, Mathiesen M, Zhou YJ, Arredondo-Peter R, Sarath G, Klucas RV. 2001. Non-symbiotic hemoglobins in rice are synthesized during germination and in differentiating cell types. *Protoplasma* **218**: 125–133.

Ross EJH, Stone JM, Elowsky CG, Arredondo-Peter R, Klucas RV, Sarath G. 2004. Activation of the *Oryza sativa* non-symbiotic haemoglobin-2 promoter by the cytokinin-regulated transcription factor, ARR1. *Journal of Experimental Botany* **55**: 1721–1731.

Rubio MC, Calvo-Begueria L, Díaz-Mendoza M, Elhiti M, Moore M, Matamoros MA, James EK, Díaz I, Pérez-Rontomé C, Villar I *et al.* 2019. Phytoglobins in the chloroplasts, nuclei and cytoplasm modulate nitric oxide and interact with abscisic acid. *The Plant Journal* **100**: 38–54.

Sainz M, Calvo-Begueria L, Pérez-Rontomé C, Wienkoop S, Abián J, Staudinger C, Bartesaghi S, Radi R, Becana M. 2015. Leghemoglobin is nitrated in functional legume nodules in a tyrosine residue within the heme cavity by a nitrite/peroxide-dependent mechanism. *The Plant Journal* **81**: 723–735.

Sainz M, Pérez-Rontomé C, Ramos J, Mulet JM, James EK, Bhattacharjee U, Petrich JW, Becana M. 2013. Plant hemoglobins may be maintained in functional form by reduced flavins in the nuclei, and confer differential tolerance to nitro-oxidative stress. *The Plant Journal* **76**: 875–887.

Sakamoto A, Sakurao SH, Fukunaga K, Matsubara T, Ueda-Hashimoto M, Tsukamoto S, Takahashi M, Morikawa H. 2004. Three distinct *Arabidopsis* hemoglobins exhibit peroxidase-like activity and differentially mediate nitrite-dependent protein nitration. *FEBS Letters* **572**: 27–32.

Sánchez C, Gates AJ, Meakin GE, Uchiumi T, Girard L, Richardson DJ, Bedmar E, Delgado MJ. 2010. Production of nitric oxide and nitrosylleghemoglobin complexes in soybean nodules in response to flooding. *Molecular Plant-Microbe Interactions* **23**: 702–711.

Sanz-Luque E, Ocaña-Calahorro F, de Montaigu A, Chamizo-Ampudia A, Llamas A, Galván A, Fernández E. 2015. THB1, a truncated hemoglobin, modulates nitric oxide levels and nitrate reductase activity. *The Plant Journal* **81**: 467–479.

Sasakura F, Uchiumi T, Shimoda Y, Suzuki A, Takenouchi K, Higashi S, Abe M. 2006. A class 1 hemoglobin gene from *Alnus firma* functions in symbiotic and nonsymbiotic tissues to detoxify nitric oxide. *Molecular Plant–Microbe Interactions* **19**: 441–450.

Scott NL, Xu Y, Shen G, Vuletich DA, Falzone CJ, Li Z, Ludwig M, Pond MP, Preimesberger MR, Bryant DA *et al.* 2011. Functional and structural characterization of the 2/2 hemoglobin from *Synechococcus* sp. PCC 7002. *Biochemistry* **49**: 7000–7011.

Seregélyes C, Mustárdy L, Ayaydin F, Sass L, Kovács L, Endre G, Lukács N, Kovács I, Vass I, Kiss GB *et al.* 2000. Nuclear localization of a hypoxia-inducible novel non-symbiotic hemoglobin in cultured alfalfa cells. *FEBS Letters* **482**: 125–130.

Shankar A, Fernandes JL, Kaur K, Sharna M, Kundu S, Pandey GK. 2018. Rice phytoglobins regulate responses under low mineral nutrients and abiotic stresses in *Arabidopsis thaliana*. *Plant, Cell & Environment* **41**: 215–230.

Shimoda Y, Nagata M, Suzuki A, Abe M, Sato S, Kato T, Tabata S, Higashi S, Uchiumi T. 2005. Symbiotic *Rhizobium* and nitric oxide induce gene expression

of non-symbiotic hemoglobin in *Lotus japonicus*. *Plant and Cell Physiology* **46**: 99–107.

Shimoda Y, Shimoda-Sasakura F, Kucho K, Kanamori N, Nagata M, Suzuki A, Abe M, Higashi S, Uchiumi T. 2009. Overexpression of class 1 plant hemoglobin genes enhances symbiotic nitrogen fixation activity between *Mesorhizobium loti* and *Lotus japonicus*. *The Plant Journal* **57**: 254–263.

Silvester WB, Berg RH, Schwintzer CR, Tjepkema JD. 2008. Oxygen response, hemoglobin, and the structure and function of vesicles. In: Pawlowski K, Newton WE, eds. *Nitrogen-fixing actinorhizal symbioses*. Dordrecht, the Netherlands: Springer, 105–146.

Smagghe BJ, Hoy JA, Percifield R, Kundu S, Hargrove MS, Sarath G, Hilbert JL, Watts RA, Dennis ES, Peacock WJ *et al.* 2009. Correlations between oxygen affinity and sequence classifications of plant hemoglobins. *Biopolymers* **91**: 1083–1096.

Sowa AW, Duff SMG, Guy PA, Hill RD. 1998. Altering hemoglobin levels changes energy status in maize cells under hypoxia. *Proceedings of the National Academy of Sciences, USA* **95**: 10317–10321.

Spyrakis F, Bruno S, Bidon-Chanal A, Luque FJ, Abbruzzetti S, Viappiani C, Dominici P, Mozzarelli A. 2011. Oxygen binding to *Arabidopsis thaliana* AHb2 nonsymbiotic hemoglobin: evidence for a role in oxygen transport. *International Union of Biochemistry and Molecular Biology Life* **63**: 355–362.

Stasolla C, Hill RD. 2017. Determining cellular responses: phytoglobins may direct the traffic. *Trends in Plant Science* **22**: 820–822.

Stasolla C, Huang S, Hill RD, Igamberdiev AU. 2019. Spatio-temporal expression of phytoglobin: a determining factor in the NO specification of cell fate. *Journal of Experimental Botany* **70**: 4365–4377.

Stewart JJ, Coyne KJ. 2011. Analysis of raphidophyte assimilatory nitrate reductase reveals unique domain architecture incorporating a 2/2 hemoglobin. *Plant Molecular Biology* **77**: 565–575.

Sturms R, DiSpirito AA, Fulton DB, Hargrove MS. 2011a. Hydroxylamine reduction to ammonium by plant and cyanobacterial hemoglobins. *Biochemistry* **50**: 10829–10835.

Sturms R, DiSpirito AA, Hargrove MS. 2011b. Plant and cyanobacterial hemoglobins reduce nitrite to nitric oxide under anoxic conditions. *Biochemistry* **50**: 3873–3878.

Sturms R, Kakar S, Trent J III, Hargrove MS. 2010. *Trema* and *Parasponia* hemoglobins reveal convergent evolution of oxygen transport in plants. *Biochemistry* **49**: 4085–4093.

Taylor ER, Nie XZ, MacGregor AW, Hill RD. 1994. A cereal haemoglobin gene is expressed in seed and root tissues under anaerobic conditions. *Plant Molecular Biology* **24**: 853–862.

Tiso M, Tejero J, Kenney C, Frizzell S, Gladwin MT. 2012. Nitrite reductase activity of nonsymbiotic hemoglobins from *Arabidopsis thaliana*. *Biochemistry* **51**: 5285–5292.

Trevaskis B, Watts RA, Andersson CR, Llewellyn DJ, Hargrove MS, Olson JS, Dennis ES, Peacock WJ. 1997. Two hemoglobin genes in *Arabidopsis thaliana*: the evolutionary origin of leghemoglobins. *Proceedings of the National Academy of Sciences, USA* **94**: 12230–12234.

Trinick MJ, Goodchild DJ, Miller C. 1989. Localization of bacteria and hemoglobin in root nodules of *Parasponia andersonii* containing both *Bradyrhizobium* strains and *Rhizobium leguminosarum* biovar *trifolii*. *Applied and Environmental Microbiology* **55**: 2046–2055.

Uchiumi T, Shimoda Y, Tsuruta T, Mukoyoshi Y, Suzuki A, Senoo K, Sato S, Kato T, Tabata S, Higashi S *et al.* 2002. Expression of symbiotic and nonsymbiotic globin genes responding to microsymbionts on *Lotus japonicus*. *Plant and Cell Physiology* **43**: 1351–1358.

Vargas P, Zardoya R. 2014. *The tree of life*. Sunderland, MA, USA: Sinauer Associates.

Vázquez-Limón C, Castro-Bustos S, Arredondo-Peter R. 2012a. Spectroscopic analysis of moss (*Ceratodon purpureus* and *Physcomitrella patens*) recombinant non-symbiotic hemoglobins. *Communicative & Integrative Biology* **5–6**: 527–530.

Vázquez-Limón C, Hoogewijs D, Vinogradov SN, Arredondo-Peter R. 2012b. The evolution of land plant hemoglobins. *Plant Science* **191–192**: 71–81.

van Velzen R, Holmer R, Bu F, Rutten L, van Zeijl A, Liu W, Santuari L, Cao Q, Sharma T, Shen D *et al.* 2018. Comparative genomics of the nonlegume *Parasponia* reveals insights into evolution of nitrogen-fixing rhizobium symbioses. *Proceedings of the National Academy of Sciences, USA* **115**: 4700–4709.

Vieweg MF, Hohnjec N, Küster H. 2005. Two genes encoding different truncated hemoglobins are regulated during root nodule and arbuscular mycorrhiza symbioses of *Medicago truncatula*. *Planta* **220**: 757–766.

Vigeolas H, Hühn D, Geigenberger P. 2011. Non-symbiotic hemoglobin-2 leads to an elevated energy state and to a combined increase in polyunsaturated fatty acids and total oil content when overexpressed in developing seeds of transgenic *Arabidopsis* plants. *Plant Physiology* **155**: 1435–1444.

Vinogradov SN, Fernández I, Hoogewijs D, Arredondo-Peter R. 2011. Phylogenetic relationships of 3/3 and 2/2 hemoglobins in Archaeplastida genomes to bacterial and other eukaryote hemoglobins. *Molecular Plant* **4**: 42–58.

Vinogradov SN, Hoogewijs DH, Bailly X, Arredondo-Peter R, Guertin M, Gough J, Dewilde S, Moens L, Vanfleteren JR. 2005. Three globin lineages belonging to two structural classes in genomes from the three kingdoms of life. *Proceedings of the National Academy of Sciences, USA* **102**: 11385–11389.

Vinogradov SN, Tinajero-Trejo M, Poole RK, Hoogewijs D. 2013. Bacterial and archaeal globins - a revised perspective. *Biochimica et Biophysica Acta* **1834**: 1789–1800.

Violante-Mota F, Tellechea E, Moran JF, Sarath G, Arredondo-Peter R. 2010. Analysis of peroxidase activity of rice (*Oryza sativa*) recombinant hemoglobin 1: implications for *in vivo* function of hexacoordinate non-symbiotic hemoglobins in plants. *Phytochemistry* **71**: 21–26.

Wang L, Rubio MC, Xin X, Zhang B, Fan Q, Wang Q, Ning G, Becana M, Duanmu D. 2019. CRISPR/Cas9 knockout of leghemoglobin genes in *Lotus japonicus* uncovers their synergistic roles in symbiotic nitrogen fixation. *New Phytologist* **224**: 818–832.

Wang X, Hargrove MS. 2013. Nitric oxide in plants: the roles of ascorbate and hemoglobin. *PLoS ONE* **8**: e82611.

Wang Y, Elhiti M, Hebelstrup KH, Hill RD, Stasolla C. 2011. Manipulation of hemoglobin expression affects *Arabidopsis* shoot organogenesis. *Plant Physiology and Biochemistry* **49**: 1108–1116.

Wang YH, Kochian LV, Doyle JJ, Garvin DF. 2003. Two tomato non-symbiotic haemoglobin genes are differentially expressed in response to diverse changes in mineral nutrient status. *Plant, Cell & Environment* **26**: 673–680.

Wany A, Gupta AK, Kumari A, Mishra S, Singh N, Pandey S, Vanvari R, Igamberdiev AU, Fernie AR, Gupta KJ. 2019. Nitrate nutrition influences multiple factors in order to increase energy efficiency under hypoxia in *Arabidopsis*. *Annals of Botany* **123**: 691–705.

Watts RA, Hunt PW, Hvitved AN, Hargrove MS, Peacock WJ, Dennnis ES. 2001. A hemoglobin from plants homologous to truncated hemoglobins from microorganisms. *Proceedings of the National Academy of Sciences, USA* **98**: 10119–10124.

Weber RE, Vinogradov SN. 2001. Nonvertebrate hemoglobins: functions and molecular adaptations. *Physiological Reviews* **81**: 569–628.

Weiland TR, Kundu S, Trent JT III, Hoy JA, Hargrove MS. 2004. Bis-histidyl hexacoordination in hemoglobins facilitates heme reduction kinetics. *Journal of the American Chemical Society* **126**: 11930–11935.

Wittenberg JB, Bolognesi M, Wittenberg BA, Guertin M. 2002. Truncated hemoglobins: a new family of hemoglobins widely distributed in bacteria, unicellular eukaryotes, and plants. *Journal of Biological Chemistry* **277**: 871–874.

Yang LX, Wang RY, Ren F, Liu J, Cheng J, Lu YT. 2005. AtGLB1 enhances the tolerance of *Arabidopsis* to hydrogen peroxide stress. *Plant and Cell Physiology* **46**: 1309–1316.

Zhang J, Buegger F, Albert A, Ghirardo A, Winkler B, Schnitzler JP, Hebelstrup KH, Durner J, Lindermayr C. 2019. Phytoglobin overexpression promotes barley growth in the presence of enhanced level of atmospheric nitric oxide. *Journal of Experimental Botany* **70**: 4521–4537.

Zhu H, Riggs AF. 1992. Yeast flavohemoglobin is an ancient protein related to globins and a reductase family. *Proceedings of the National Academy of Sciences, USA* **89**: 5015–5019.

See also the Commentary on this article by Catalán & Vogel, **227**: 1587–1590.